ESTIMATION OF COVARIANCE MATRICES WITH LINEAR STRUCTURE AND MOVING AVERAGE PROCESSES OF FINITE ORDER

by

T. W. ANDERSON
Stanford University

TECHNICAL REPORT NO. 6

OCTOBER 29, 1971

PREPARED UNDER THE AUSPICES OF

OFFICE OF NAVAL RESEARCH CONTRACT #N00014-67-A-0112-0030

T. W. ANDERSON, PROJECT DIRECTOR

DEPARTMENT OF STATISTICS

STANFORD UNIVERSITY

STANFORD, CALIFORNIA

[illegible] 1, 1971

ERRATA TO

ESTIMATION OF COVARIANCE MATRICES WITH LINEAR STRUCTURE AND MOVING AVERAGE PROCESSES OF FINITE ORDER

by

T. W. Anderson

Technical Report No. 6
October 29, 1971

	ℓ 9	Add ")" after "in general"
	ℓ 4	"g_n" at end of line should be "g_m"
	ℓ 14	Add "," after "nonsingular"
	ℓ 4	Delete ","
	ℓ 3	(1.26) should be (1.29)
	ℓ 2	Period should be a comma
	ℓ 6	"elmination" should be "elimination"
	(3.29)	4(p+1) should be p+1 4p should be 2(p-1)
	(3.38)	$\frac{1}{\sigma^4(1-\alpha^2)^3}$ should be $\frac{1+\alpha^2}{\sigma^4(1-\alpha^2)^3}$
	(3.42)	Add new paragraph "Let" before (3.42) $\sum_{j=1}^{T-h}$ should be $\sum_{j=1}^{p-h}$ Add , h=0, 1, ... , p-1.
[illegible]	(3.48)	"$-$" should be "$+$"
[illegible]	ℓ 4	"k_h" should be "k_m"
[illegible]	ℓ 1	$\sum_{j=1}^{p}$ should be $\sum_{j=1}^{p}$ Change y to z two times
32	(3.59) ℓ 3	$(1-z_g z_h)/[1-(z_g z_h)^p]$ should be $[1-(z_g z_h)^p]/(1-z_g z_h)$
32	ℓ 9	Add comma before "for"
32	ℓ 1	$(2\pi)^2$ should be $(2\pi)^3$
34 and 35		Should be interchanged

ERRATA
Page 2

pg. 33	(3.66)	$z_g z_h$ should be $y_g y_h$ two times
pg. 34	(3.59)	$\frac{1}{(2\pi)^2}$ should be $\frac{2}{(2\pi)^3}$
pg. 34	(3.60)	$\frac{1}{(2\pi)^2}$ should be $\frac{4}{(2\pi)^3}$
pg. 35	(3.73) ℓ 3	$\sum_{k,\ =1}^{p}$ should be $\sum_{k,\ell=1}^{p}$
		$g_{k_\ell}^{(g)}\ b_{q_\ell}$ should be $g_{k\ell}^{(g)}\ b_{q\ell}$
pg. 35	(3.73) ℓ 4	$g_{k}^{(g)}\ b_{q}$ should be $g_{k\ell}^{(g)}\ b_{q\ell}$
pg. 35	(3.73) ℓ 5	Comma between $j-k'$ and $q-\ell'$
pg. 36	(3.79) ℓ 3	$\frac{1}{2\pi(p-K)}$ should be $\frac{1}{(2\pi)^2(p-K)}$
pg. 36	(3.79) ℓ 4	Add $\sum_{g=0}^{m} y_g$ after $b_{\ell'}$, $e^{i\lambda\ell'}$
pg. 36	(3.79) ℓ 5	$\frac{1}{(2\pi)(p-K)}$ should be $\frac{1}{(2\pi)^2(p-K)}$
pg. 37	(3.81) ℓ 1	$\frac{1}{(2\pi)^2}$ should be $\frac{4}{(2\pi)^3}$
pg. 37	(3.83) ℓ 3 and 4	$\frac{4}{(2\pi)^2}$ should be $\frac{4}{(2\pi)^3}$
pg. 38	(3.85) ℓ 2	$\frac{4}{2\pi}$ should be $\frac{4}{(2\pi)^3}$
pg. 38	(3.96) ℓ 2	$\frac{4}{(2\pi)^2}$ should be $\frac{4}{(2\pi)^3}$
pg. 40	(4.6) ℓ 1	$(\underset{\sim}{\hat{\beta}}{}^*_N - \underset{\sim}{\hat{\beta}})$ should be $(\underset{\sim}{\hat{\beta}}{}^*_N - \underset{\sim}{\hat{\beta}}{}_N)$
pg. 41	ℓ 4	$\underset{\sim}{C}$ should be $\underset{\sim}{c}$
pg. 45	(5.19) ℓ 1	d_n should be d_h
pg. 47	(5.25)	$g=1, \ldots, m.$ should be $g=0, 1, \ldots, m.$
pg. 47	ℓ 4	Add $(n \times n)$ after $\underset{\sim}{C}^*$

ADDITIONAL REFERENCES

Box, George E. P., and Gwilym M. Jenkins (1970), Time Series Analysis Forecasting and Control, Holden-Day Inc., San Francisco.

Clevenson, M. Lawrence (1970), "Asymptotically efficient estimates of the parameters of a moving average time series", Stanford University, Statistics Department.

Hannan, E. J. (1969), "The estimation of mixed moving average autoregressive systems", Biometrika, **56**, pp. 579-594.

Walker, A. M. (1962), "Large sample estimation of parameters for autoregressive processes with moving-average residuals", Biometrika, **49**, pp. 117-131.

1. Introduction and Review of Earlier Work.

This paper deals with estimation problems in which one or more observations are made on a p-component vector $\underset{\sim}{X}$ with mean vector $\mathcal{E}\underset{\sim}{X} = \underset{\sim}{\mu}$ and covariance matrix

$$(1.1) \qquad \mathcal{C}(\underset{\sim}{X}) = \mathcal{E}(\underset{\sim}{X}-\underset{\sim}{\mu})(\underset{\sim}{X}-\underset{\sim}{\mu})' = \underset{\sim}{\Sigma} ,$$

where $\underset{\sim}{\mu}$ and $\underset{\sim}{\Sigma}$ may have linear structure. The mean $\underset{\sim}{\mu}$ may be a linear combination

$$(1.2) \qquad \underset{\sim}{\mu} = \sum_{j=1}^{r} \beta_j \underset{\sim}{z}_j$$

of known p-component vectors, $\underset{\sim}{z}_1, \ldots , \underset{\sim}{z}_r$, which are assumed (for convenience) to be linearly independent. The covariance matrix may be a linear combination

$$(1.3) \qquad \underset{\sim}{\Sigma} = \sum_{g=0}^{m} \sigma_g \underset{\sim}{G}_g$$

of known symmetric $p \times p$ matrices $\underset{\sim}{G}_0, \underset{\sim}{G}_1, \ldots , \underset{\sim}{G}_m$, which are assumed to be linearly independent; it is also assumed that there is at least one set $\sigma_0, \sigma_1, \ldots , \sigma_m$ such that (1.3) is positive definite. The coefficients $\beta_1, \ldots , \beta_r$ and $\sigma_0, \sigma_1, \ldots , \sigma_m$ are parameters.

If $\underset{\sim}{\Sigma}$ is known or known to within a constant of proportionality and one observation $\underset{\sim}{x}$ is made on $\underset{\sim}{X}$, the model is the familiar one of regression analysis. The best linear unbiased estimates or Markov estimates of $\beta_1, \ldots , \beta_r$ are the solutions to the normal equations

$$(1.4) \qquad \sum_{i=1}^{r} \underset{\sim}{z}_j' \underset{\sim}{\Sigma}^{-1} \underset{\sim}{z}_i \hat{\beta}_i = \underset{\sim}{z}_i' \underset{\sim}{\Sigma}^{-1} \underset{\sim}{x} , \qquad j=1, \ldots , r .$$

If $\underset{\sim}{X}$ has a normal distribution, (1.4) are the likelihood equations, obtained by setting equal to 0 the derivatives of the likelihood function with respect to $\beta_1, \ldots , \beta_r$, and the solution constitutes the maximum likelihood estimates. In any case the estimates are unbiased, $\mathcal{E} \hat{\beta}_i = \beta_i$, $i=1, \ldots , r$, and the covariance matrix of the estimates is

$$(1.5) \qquad [\mathcal{C}(\hat{\beta}_i, \hat{\beta}_j)] = [\underset{\sim}{z}_i' \underset{\sim}{\Sigma}^{-1} \underset{\sim}{z}_j]^{-1} .$$

If there are N observations on $\underset{\sim}{X}$, say $\underset{\sim}{x}_1, \ldots , \underset{\sim}{x}_N$, the best linear unbiased estimates and maximum likelihood estimates under normality are the solution to (1.4) with $\underset{\sim}{x}$ replaced by the sample mean

$$(1.6) \qquad \bar{\underset{\sim}{x}} = \frac{1}{N} \sum_{\alpha=1}^{N} \underset{\sim}{x}_\alpha ,$$

and the covariance matrix of the estimates is $1/N$ times (1.5).

Estimation of $\sigma_0, \sigma_1, \ldots , \sigma_m$ was considered by T. W. Anderson* (1969), (1970) when several observations were made on $\underset{\sim}{X}$ and $\underset{\sim}{\mu}$ was completely unspecified. In this present review we shall assume initially that $\underset{\sim}{\mu}$ is known and suitably modify the statements of the earlier papers. We assume that $\underset{\sim}{X}$ has a normal distribution and that there are N observations $\underset{\sim}{x}_1, \ldots , \underset{\sim}{x}_N$ on $\underset{\sim}{X}$. (N is not necessarily as large as p ; in fact, N may be 1 in some cases.) Let

---

* The 1970 paper was written first, but there was a delay of four years between its receipt by the editors and its publication.

$$C = \frac{1}{N} \sum_{\alpha=1}^{N} (x_\alpha - \mu)(x_\alpha - \mu)' . \tag{1.7}$$

Then maximum likelihood estimates of $\sigma_0, \sigma_1, \ldots, \sigma_m$ are a solution of the likelihood equations

$$\operatorname{tr}\left(\sum_{h=0}^{m} \hat{\sigma}_h G_h\right)^{-1} G_g = \operatorname{tr}\left(\sum_{h=0}^{m} \hat{\sigma}_h G_h\right)^{-1} G_g \left(\sum_{h=0}^{m} \hat{\sigma}_h G_h\right)^{-1} C , \quad g=0, 1, \ldots, m ; \tag{1.8}$$

these equations result from setting equal to 0 the derivatives of the likelihood function with respect to $\sigma_0, \sigma_1, \ldots, \sigma_m$. There is at least one solution $\hat{\sigma}_0, \hat{\sigma}_1, \ldots, \hat{\sigma}_m$ to (1.8) such that

$$\hat{\Sigma} = \sum_{g=0}^{m} \hat{\sigma}_g G_g \tag{1.9}$$

is positive definite. (The argument given by T. W. Anderson in (1970) was stated for C positive definite, but that assumption is not needed in general. If there is more than one solution to the likelihood equations, the absolute maximum to the likelihood function is attained by the solution minimizing $|\hat{\Sigma}|$. The estimates $\hat{\sigma}_0, \hat{\sigma}_1, \ldots, \hat{\sigma}_m$ are consistent and asymptotically efficient as $N \to \infty$; $\sqrt{N}(\hat{\sigma}_0 - \sigma_0), \sqrt{N}(\hat{\sigma}_1 - \sigma_1), \ldots, \sqrt{N}(\hat{\sigma}_m - \sigma_m)$ have a limiting normal distribution with means 0 and covariance matrix

$$[\tfrac{1}{2} \operatorname{tr} \Sigma^{-1} G_h \Sigma^{-1} G_g]^{-1} . \tag{1.10}$$

These results follow from the usual asymptotic theory of maximum likelihood estimates.

Now let us consider the estimation problem when both $\underset{\sim}{\mu}$ given by (1.2) and $\underset{\sim}{\Sigma}$ given by (1.3) are unknown. We are to estimate $\beta_1, \ldots, \beta_r$ and $\sigma_0, \sigma_1, \ldots, \sigma_m$. When $\underset{\sim}{X}$ is normally distributed and $\underset{\sim}{x}_1, \ldots, \underset{\sim}{x}_N$ are observed, the likelihood equations are

$$(1.11) \qquad \sum_{i=1}^{r} \underset{\sim}{z}_j' \hat{\underset{\sim}{\Sigma}}^{-1} \underset{\sim}{z}_i \hat{\beta}_i = \underset{\sim}{z}_j' \hat{\underset{\sim}{\Sigma}}^{-1} \bar{\underset{\sim}{x}} , \qquad j=1, \ldots, r ,$$

$$(1.12) \qquad \operatorname{tr}\left(\sum_{h=0}^{m} \hat{\sigma}_h \underset{\sim}{G}_h\right)^{-1} \underset{\sim}{G}_g = \operatorname{tr}\left(\sum_{h=0}^{m} \hat{\sigma}_h \underset{\sim}{G}_h\right)^{-1} \underset{\sim}{G}_g \left(\sum_{h=0}^{m} \hat{\sigma}_h \underset{\sim}{G}_h\right)^{-1} \hat{\underset{\sim}{C}} , \qquad g=0, 1, \ldots, m ,$$

where

$$(1.13) \qquad \hat{\underset{\sim}{C}} = \frac{1}{N} \sum_{\alpha=1}^{N} (\underset{\sim}{x}_\alpha - \hat{\underset{\sim}{\mu}})(\underset{\sim}{x}_\alpha - \hat{\underset{\sim}{\mu}})' ,$$

$$(1.14) \qquad \hat{\underset{\sim}{\mu}} = \sum_{j=1}^{r} \hat{\beta}_j \underset{\sim}{z}_j ,$$

and $\hat{\underset{\sim}{\Sigma}}$ is given by (1.9). Then the estimates are consistent and asymptotically efficient as $N \to \infty$. Moreover, $\sqrt{N}(\hat{\beta}_1 - \beta_1), \ldots, \sqrt{N}(\hat{\beta}_r - \beta_r)$ and $\sqrt{N}(\hat{\sigma}_0 - \sigma_0), \sqrt{N}(\hat{\sigma}_1 - \sigma_1), \ldots, \sqrt{N}(\hat{\sigma}_m - \sigma_m)$ have a limiting normal distribution in which the two sets are independent and each set has the covariance matrix given previously, (1.5) and (1.10). (We note in passing that asymptotically as $N \to \infty$, $\underset{\sim}{C}$, $\hat{\underset{\sim}{C}}$, and $(1/N) \sum_{\alpha=1}^{N} (\underset{\sim}{x}_\alpha - \bar{\underset{\sim}{x}})(\underset{\sim}{x}_\alpha - \bar{\underset{\sim}{x}})'$ are equivalent; replacement of $\hat{\underset{\sim}{C}}$ in (1.12) by the last matrix above represents a simplification of the equations.)

The relation between the estimation of $\beta_1, \ldots, \beta_r$ and $\sigma_0, \sigma_1, \ldots, \sigma_m$ separately was indicated by T. W. Anderson (1969).

Consider (1.4) for $\underset{\sim}{\Sigma}$ known and consider (1.8) for $\underset{\sim}{\mu}$ known. In the case of normality the covariance between the i,j-th and k,ℓ-th elements of $\underset{\sim}{C}$ is

$$(1.15) \qquad \mathrm{Cov}(c_{ij}, c_{k\ell}) = \frac{1}{N}(\sigma_{ik}\,\sigma_{j\ell} + \sigma_{i\ell}\,\sigma_{jk}) .$$

Let

$$(1.16) \qquad \underset{\sim}{c} = \begin{pmatrix} c_{11} \\ c_{22} \\ \vdots \\ c_{pp} \\ c_{12} \\ \vdots \\ c_{p-1,\,p} \end{pmatrix}, \quad \underset{\sim}{\sigma} = \begin{pmatrix} \sigma_{11} \\ \sigma_{22} \\ \vdots \\ \sigma_{pp} \\ \sigma_{12} \\ \vdots \\ \sigma_{p-1,\,p} \end{pmatrix}, \quad \underset{\sim}{g}_h = \begin{pmatrix} g_{11}^{(h)} \\ g_{22}^{(h)} \\ \vdots \\ g_{pp}^{(h)} \\ g_{12}^{(h)} \\ \vdots \\ g_{p-1,\,p}^{(h)} \end{pmatrix},$$

where $\underset{\sim}{G}_h = \left(g_{ij}^{(h)}\right)$. Then $\mathcal{E}\,\underset{\sim}{c} = \underset{\sim}{\sigma}$ and (1.3) can be written as

$$(1.17) \qquad \underset{\sim}{\sigma} = \sum_{h=0}^{m} \sigma_h\, \underset{\sim}{g}_h ,$$

which is of the form (1.2) with $\underset{\sim}{\mu}$ replaced by $\underset{\sim}{\sigma}$, $\beta_1, \ldots, \beta_r$ replaced by $\sigma_0, \sigma_1, \ldots, \sigma_m$, $\underset{\sim}{z}_1, \ldots, \underset{\sim}{z}_r$ replaced by $\underset{\sim}{g}_0, \underset{\sim}{g}_1, \ldots, \underset{\sim}{g}_n$, and $\mathcal{E}\underset{\sim}{X} = \underset{\sim}{\mu}$ by $\mathcal{E}\underset{\sim}{c} = \underset{\sim}{\sigma}$. Then (1.15) can be written as

$$(1.18) \qquad \mathcal{C}(\underset{\sim}{c}) = \mathcal{E}(\underset{\sim}{c}-\underset{\sim}{\sigma})(\underset{\sim}{c}-\underset{\sim}{\sigma})' = \underset{\sim}{\Phi} ,$$

where $\underset{\sim}{\Phi} = (\phi_{ij,k\ell})$ for $i \leq j$, $k \leq \ell$ and

(1.19) $$\phi_{ij,k\ell} = \frac{1}{N}(\sigma_{ik}\,\sigma_{j\ell} + \sigma_{i\ell}\,\sigma_{jk}) , \qquad i \leq j,\ k \leq \ell ;$$

(1.18) is of the form (1.1) with $\underset{\sim}{X}$ replaced by $\underset{\sim}{c}$, $\underset{\sim}{\mu}$ replaced by $\underset{\sim}{\sigma}$, and $\underset{\sim}{\Sigma}$ replaced by $\underset{\sim}{\Phi}$.

It was shown by T. W. Anderson (1969) that

(1.20) $$\underset{\sim}{g}_h' \, \underset{\sim}{\Phi}^{-1} \, \underset{\sim}{c} = \tfrac{1}{2} \operatorname{tr} \underset{\sim}{\Sigma}^{-1} \, \underset{\sim}{G}_h \, \underset{\sim}{\Sigma}^{-1} \, \underset{\sim}{C} ;$$

that is, the bilinear form on the left-hand side of (1.20) is algebraically identical to the right-hand side, which is a form appearing in the likelihood equations for $\hat{\sigma}_0, \hat{\sigma}_1, \ldots , \hat{\sigma}_m$. Substitution of $\underset{\sim}{g}_g$ for $\underset{\sim}{c}$ in (1.20) yields

(1.21) $$\underset{\sim}{g}_h' \, \underset{\sim}{\Phi}^{-1} \, \underset{\sim}{g}_g = \tfrac{1}{2} \operatorname{tr} \underset{\sim}{\Sigma}^{-1} \, \underset{\sim}{G}_h \, \underset{\sim}{\Sigma}^{-1} \, \underset{\sim}{G}_g .$$

Thus the "normal equations"

(1.22) $$\sum_{f=0}^{m} \underset{\sim}{g}_h' \, \underset{\sim}{\Phi}^{-1} \, \underset{\sim}{g}_f \, \hat{\sigma}_f = \underset{\sim}{g}_h' \, \underset{\sim}{\Phi}^{-1} \, \underset{\sim}{c} , \qquad h=0, 1, \ldots , m ,$$

are identical to the equations

(1.23) $$\sum_{f=0}^{m} \hat{\sigma}_f \operatorname{tr} \underset{\sim}{\Sigma}^{-1} \, \underset{\sim}{G}_g \, \underset{\sim}{\Sigma}^{-1} \, \underset{\sim}{G}_f = \operatorname{tr} \underset{\sim}{\Sigma}^{-1} \, \underset{\sim}{G}_g \, \underset{\sim}{\Sigma}^{-1} \, \underset{\sim}{C} , \qquad g=0, 1, \ldots , m ,$$

If the equations (1.23) were available, they would give "estimates" which were linear in $\underset{\sim}{C}$ and unbiased, and among such "estimates" they would have minimum variance. (Since $\underset{\sim}{C}$ is sufficient for $\underset{\sim}{\Sigma}$, these would be minimum variance unbiased "estimates" of $\sigma_0, \sigma_1, \ldots , \sigma_m$.)

The covariance matrix of the "estimates" would be $1/N$ times (1.10). However, these "estimates" are unavailable since $\underset{\sim}{\Sigma}$ is unknown; in fact, the problem is to estimate $\underset{\sim}{\Sigma}$.

The likelihood equations (1.8) for $\hat{\sigma}_0, \hat{\sigma}_1, \ldots, \hat{\sigma}_m$ can be written

$$(1.24) \qquad \sum_{f=0}^{m} \operatorname{tr} \hat{\underset{\sim}{\Sigma}}^{-1} \underset{\sim}{G}_g \hat{\underset{\sim}{\Sigma}}^{-1} \underset{\sim}{G}_f \hat{\sigma}_f = \operatorname{tr} \hat{\underset{\sim}{\Sigma}}^{-1} \underset{\sim}{G}_g \hat{\underset{\sim}{\Sigma}}^{-1} \underset{\sim}{C}, \quad g=0, 1, \ldots, m,$$

by multiplying the left-hand side of (1.8) by $\underset{\sim}{I} = \sum_{f=0}^{m} \hat{\sigma}_f \underset{\sim}{G}_f \hat{\underset{\sim}{\Sigma}}^{-1}$. These equations are similar to (1.23), but $\underset{\sim}{\Sigma}$ in (1.23) has been replaced by $\hat{\underset{\sim}{\Sigma}}$. As will be shown later, the form (1.23) suggests computational procedures and asymptotic properties.

One of the probability models in which the covariance matrix has the form (1.3) is a moving average stationary stochastic process of finite order. Let

$$(1.25) \qquad x_t = \sum_{g=0}^{m} \alpha_g v_{t-g}, \qquad t = \ldots, -1, 0, 1, \ldots,$$

where $\mathcal{E} v_t = 0$, $\mathcal{E} v_t^2 = \sigma^2$, and $\mathcal{E} v_t v_s = 0$, $t \neq s$. Then $\mathcal{E} x_t = 0$ and

$$(1.26) \qquad \begin{aligned} \sigma_h = \sigma_{-h} &= \mathcal{E} x_t x_{t+h} \\ &= \sigma^2 \sum_{g=0}^{m-h} \alpha_g \alpha_{g+h}, \qquad h=0, 1, \ldots, m, \\ &= 0, \qquad h = m+1, \ldots, \end{aligned}$$

The vector $\underset{\sim}{x} = (x_1, \ldots, x_p)'$ has the covariance matrix (1.3) with $\underset{\sim}{G}_0 = \underset{\sim}{I}$,

$$(1.27) \qquad \underset{\sim}{G}_1 = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 1 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix},$$

$$(1.28) \qquad \underset{\sim}{G}_2 = \begin{pmatrix} 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix},$$

etc. Then

$$(1.29) \quad \underset{\sim}{\Sigma} = \begin{pmatrix} \sigma_0 & \sigma_1 & \cdots & \sigma_m & 0 & \cdots & 0 \\ \sigma_1 & \sigma_0 & \cdots & \sigma_{m-1} & \sigma_m & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \sigma_m & \sigma_{m-1} & \cdots & \sigma_0 & \sigma_1 & \cdots & 0 \\ 0 & \sigma_m & \cdots & \sigma_1 & \sigma_0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & \sigma_0 \end{pmatrix}.$$

When $\sigma_0, \sigma_1, \ldots, \sigma_m$ are defined by (1.26) for real $\alpha_0, \alpha_1, \ldots, \alpha_m$, $\sigma^2 > 0$, $\underset{\sim}{\Sigma}$ given by (1.19) is positive definite (of any order).

If $\underset{\sim}{x} = (x_1, \ldots, x_p)'$ is observed, the maximum likelihood estimates of $\sigma_0, \sigma_1, \ldots, \sigma_m$ are defined by (1.8) or (1.24) where $\underset{\sim}{C} = \underset{\sim}{x}\underset{\sim}{x}'$ (for $\underset{\sim}{\mu} = \underset{\sim}{0}$ and $N = 1$). Then the right-hand side of (1.8) and (1.24) is

$$\text{tr}\, \hat{\underset{\sim}{\Sigma}}^{-1} \underset{\sim}{G}_g \hat{\underset{\sim}{\Sigma}}^{-1} \underset{\sim}{x}\underset{\sim}{x}' = \underset{\sim}{x}' \hat{\underset{\sim}{\Sigma}}^{-1} \underset{\sim}{G}_g \hat{\underset{\sim}{\Sigma}}^{-1} \underset{\sim}{x} . \tag{1.30}$$

In the case of the moving average model one may be interested in the parameters $\alpha_1, \ldots, \alpha_m, \sigma^2$, with $\alpha_0 = 1$. If the estimates $\hat{\sigma}_0, \hat{\sigma}_1, \ldots, \hat{\sigma}_m$ are such that the estimated covariance matrix obtained by replacing $\sigma_0, \sigma_1, \ldots, \sigma_m$ in (1.29) by $\hat{\sigma}_0, \hat{\sigma}_1, \ldots, \hat{\sigma}_m$ is positive definite for (1.29) of every order, then the estimate of the spectral density

$$f(\lambda) = \frac{1}{2\pi} \sum_{g=-m}^{m} \sigma_g \cos \lambda g \tag{1.31}$$

is positive and the equations (1.26) with $\sigma_0, \sigma_1, \ldots, \sigma_m$ replaced by $\hat{\sigma}_0, \hat{\sigma}_1, \ldots, \hat{\sigma}_m$ and $\alpha_1, \ldots, \alpha_m$ and σ^2 replaced by $\hat{\alpha}_1, \ldots, \hat{\alpha}_m$, $\hat{\sigma}^2$ can be solved for $\hat{\alpha}_1, \ldots, \hat{\alpha}_m, \hat{\sigma}^2 > 0$. The polynomial equation associated with the moving average (1.25) is

$$\sum_{g=0}^{m} \alpha_g z^{m-g} = 0 . \tag{1.32}$$

If the roots are required to be not greater than one in absolute value, $\alpha_1, \ldots, \alpha_m$ and σ^2 are uniquely determined by $\sigma_0, \sigma_1, \ldots, \sigma_m$. [See Section 5.7 of T. W. Anderson (1971).]

In this case of the time series problem where $N = 1$, we are interested in the asymptotic theory when $p \to \infty$. These properties will be studied later.

It was pointed out by T. W. Anderson (1969), (1970) that the model described above is appropriate for many problems of the analysis of variance. For example, let

$$x_{i\alpha} = \bar{\mu} + (\mu_i - \bar{\mu}) + u_\alpha + v_{i\alpha}, \qquad i=1, \ldots, p, \quad \alpha=1, \ldots, N, \tag{1.33}$$

where $\bar{\mu} = \sum_{i=1}^{p} \mu_i / p$, $\mathcal{E} u_\alpha = 0$, $\mathcal{E} v_{i\alpha} = 0$, $\mathcal{E} u_\alpha^2 = \sigma_u^2$, $\mathcal{E} v_{i\alpha}^2 = \sigma_v^2$, and all u_α's and $v_{i\alpha}$'s independent. Then $\mathcal{E} x_{i\alpha} = \mu_i$ and the covariance matrix of $x_{1\alpha}, \ldots, x_{p\alpha}$ is given by (1.3) with $\underset{\sim}{G}_0 = \underset{\sim}{I}$ and $\underset{\sim}{G}_1 = \underset{\sim}{\varepsilon}\underset{\sim}{\varepsilon}'$, where $\underset{\sim}{\varepsilon} = (1, \ldots, 1)'$. This is a mixed model in the analysis of variance. The overall mean is $\underset{\sim}{\mu}$, the $(\mu_i - \bar{\mu})$'s are the fixed factor effects and the u_α's are the random factor effects.

The factor analysis model when the factor loadings are known was studied by T. W. Anderson (1970). This model is particularly appropriate for one form of Guttman's quasi-simplex.

Hartley and Rao (1967) have given the derivative equations (1.4) and (1.8) for $N = 1$ when $\underset{\sim}{G}_0, \underset{\sim}{G}_1, \ldots, \underset{\sim}{G}_m$ are generated by models of the analysis of variance. Their proposals for solution are different from the one presented by T. W. Anderson (1970) and the one presented in Section 2 of this paper. Other references were given by T. W. Anderson (1969), (1970).

2. Computation of Estimates for Covariance Matrices

The set of equations (1.24) suggests an iterative method of solving the likelihood equations for $\hat{\sigma}_0, \hat{\sigma}_1, \ldots, \hat{\sigma}_m$. Let $\hat{\sigma}_0^{(0)}, \hat{\sigma}_1^{(0)}, \ldots, \hat{\sigma}_m^{(0)}$ be an initial set of values; these may be values given a priori or they may be estimates obtained in another way. Let $\hat{\sigma}_0^{(i)}, \hat{\sigma}_1^{(i)}, \ldots, \hat{\sigma}_m^{(i)}$ be the solutions to

$$(2.1) \quad \sum_{f=0}^{m} \operatorname{tr} \hat{\Sigma}_{i-1}^{-1} G_g \hat{\Sigma}_{i-1}^{-1} G_f \hat{\sigma}_f = \operatorname{tr} \hat{\Sigma}_{i-1}^{-1} G_g \hat{\Sigma}_{i-1}^{-1} C, \quad g=0, 1, \ldots, m, \quad i=1, 2, \ldots,$$

where

$$(2.2) \quad \hat{\Sigma}_{i-1} = \sum_{h=0}^{m} \hat{\sigma}_h^{(i-1)} G_h, \qquad i=1, 2, \ldots .$$

The equations (2.1) can also be written

$$(2.3) \quad \sum_{f=0}^{m} g_g' \hat{\Phi}_{i-1}^{-1} g_f \hat{\sigma}_f = g_g' \hat{\Phi}_{i-1}^{-1} c, \qquad g=0, 1, \ldots, m,$$

where $\hat{\Phi}_{i-1}$ is formed from $\hat{\Sigma}_{i-1}$ as Φ is formed from Σ.

Lemma 2.1. *If $\sum_{h=0}^{m} \sigma_h G_h$ is nonsingular*

$$(2.4) \quad \left[\operatorname{tr} \left(\sum_{h=0}^{m} \sigma_h G_h\right)^{-1} G_g \left(\sum_{h=0}^{m} \sigma_h G_h\right)^{-1} G_f\right]$$

is positive definite.

Proof. For $(y_0, y_1, \ldots, y_m) \neq (0, 0, \ldots, 0)$

$$(2.5) \quad \sum_{g,f=0}^{m} \operatorname{tr} \left(\sum_{h=0}^{m} \sigma_h G_h\right)^{-1} G_g \left(\sum_{h=0}^{m} \sigma_h G_h\right)^{-1} G_f y_g y_f$$

$$= \operatorname{tr} \left[\left(\sum_{h=0}^{m} \sigma_h G_h\right)^{-1} \sum_{g=0}^{m} y_g G_g\right] \left[\left(\sum_{h=0}^{m} \sigma_h G_h\right)^{-1} \sum_{f=0}^{m} y_f G_f\right]'$$

$$> 0$$

because $\underset{\sim}{G}_0, \underset{\sim}{G}_1, \ldots, \underset{\sim}{G}_m$ are linearly independent. Q.E.D.

If $\hat{\underset{\sim}{\Sigma}}_{i-1}$ is positive definite, the matrix of the coefficients of $\hat{\sigma}_0, \hat{\sigma}_1, \ldots, \hat{\sigma}_m$ on the left-hand side of (2.1) is positive definite, and hence there is a unique solution. (The iterative procedure suggested here can be considered as an approximation to the procedure proposed by T. W. Anderson (1970) on page 6; the present proposal is computationally simpler and its properties can be studied more easily.) The iteration may be stopped at the i-th stage if $\hat{\sigma}_0^{(i)}, \hat{\sigma}_1^{(i)}, \ldots, \hat{\sigma}_m^{(i)}$ does not differ by much from $\hat{\sigma}_0^{(i-1)}, \hat{\sigma}_1^{(i-1)}, \ldots, \hat{\sigma}_m^{(i-1)}$.

Since $\mathcal{E}\underset{\sim}{C} = \underset{\sim}{\Sigma}$, given by (1.3), unbiased estimates of $\sigma_0, \sigma_1, \ldots, \sigma_m$ can be obtained as the solutions to

$$(2.6) \qquad \sum_{f=0}^{m} \hat{\sigma}_f \operatorname{tr} \underset{\sim}{\Theta}\,\underset{\sim}{G}_g\,\underset{\sim}{\Theta}\,\underset{\sim}{G}_f = \operatorname{tr} \underset{\sim}{\Theta}\,\underset{\sim}{G}_g\,\underset{\sim}{\Theta}\,\underset{\sim}{C}, \qquad g=0, 1, \ldots, m,$$

for an arbitrary positive definite matrix $\underset{\sim}{\Theta}$. These estimates (under normality) have covariances

$$(2.7) \qquad \mathcal{E}(\hat{\sigma}_f, \hat{\sigma}_g) = \frac{2}{N} \sum_{h,j,k,\ell=0}^{m} m^{fh}\, m^{gj}\, \sigma_k\, \sigma_\ell \operatorname{tr} \underset{\sim}{\Theta}\,\underset{\sim}{G}_h\,\underset{\sim}{\Theta}\,\underset{\sim}{G}_\ell\,\underset{\sim}{\Theta}\,\underset{\sim}{G}_j\,\underset{\sim}{\Theta}\,\underset{\sim}{G}_k, \qquad f, g=0, 1, \ldots, m,$$

where $(m^{fh}) = (m_{fh})^{-1}$ and

$$(2.8) \qquad m_{fh} = \operatorname{tr} \underset{\sim}{\Theta}\,\underset{\sim}{G}_f\,\underset{\sim}{\Theta}\,\underset{\sim}{G}_h, \qquad f, h=0, 1, \ldots, m.$$

As $N \to \infty$, these estimates are consistent and $\sqrt{N}(\hat{\sigma}_0-\sigma_0), \sqrt{N}(\hat{\sigma}_1-\sigma_1), \ldots, \sqrt{N}(\hat{\sigma}_m-\sigma_m)$ have a limiting normal distribution. [If $\underset{\sim}{\mu}$ is unknown, $\underset{\sim}{C}$ in (2.6) could be replaced by $[N/(N-1)]\hat{\underset{\sim}{C}}$.]

The equations (2.6) can be obtained from (1.23) by replacing $\underset{\sim}{\Sigma}^{-1}$ by $\underset{\sim}{\Theta}$. A particular choice of $\underset{\sim}{\Theta}$ is $\underset{\sim}{I}$; this substitution corresponds to the Markov estimates when the variances of $\sqrt{N}\, c_{ii}$ are proportional to 2 , the variances of $\sqrt{N}\, c_{ij}$, $i \neq j$, are proportional to 1 and every covariance is 0 .

To obtain asymptotically efficient estimates of $\sigma_0, \sigma_1, \ldots, \sigma_m$ only one step in the iteration is needed if the initial estimates are consistent. See Section 4. If $N > 1$ and $\underset{\sim}{\mu}$ is unknown, $\underset{\sim}{C}$ may be replaced by $(1/N) \sum_{\alpha=1}^{N} (\underset{\sim}{x}_\alpha - \bar{\underset{\sim}{x}})(\underset{\sim}{x}_\alpha - \bar{\underset{\sim}{x}})'$ in the computation; one may wish to multiply the solution by the factor $N/(N-1)$.

The solution of (2.1) requires evaluation of quantities such as

$$\mathrm{tr}\, \underset{\sim}{A}^{-1} \underset{\sim}{B}\, \underset{\sim}{A}^{-1} \underset{\sim}{L} , \tag{2.9}$$

where $\underset{\sim}{A}$, $\underset{\sim}{B}$, and $\underset{\sim}{L}$ are symmetric and $\underset{\sim}{A}$ is positive definite. Finding $\underset{\sim}{A}^{-1}$ corresponds to solving

$$\underset{\sim}{A}\underset{\sim}{X} = \underset{\sim}{I} . \tag{2.10}$$

The "forward solution" of a method of pivotal condensation or successive elimination corresponds to multiplying (2.10) on the left by a triangular matrix $\underset{\sim}{F}$ to obtain

$$\underset{\sim}{T}\underset{\sim}{X} = \underset{\sim}{F} , \tag{2.11}$$

where

$$(2.12) \qquad \underset{\sim}{F} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ f_{21} & 1 & 0 & \cdots & 0 \\ f_{31} & f_{32} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ f_{p1} & f_{p2} & f_{p3} & \cdots & 1 \end{pmatrix},$$

$$(2.13) \qquad \underset{\sim}{T} = \underset{\sim\sim}{FA} = \begin{pmatrix} t_{11} & t_{12} & t_{13} & \cdots & t_{1p} \\ 0 & t_{22} & t_{23} & \cdots & t_{2p} \\ 0 & 0 & t_{32} & \cdots & t_{3p} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & t_{pp} \end{pmatrix}.$$

Then

$$(2.14) \qquad \underset{\sim\sim\sim}{FAF'} = \underset{\sim\sim}{TF'} = \begin{pmatrix} t_{11} & 0 & 0 & \cdots & 0 \\ 0 & t_{22} & 0 & \cdots & 0 \\ 0 & 0 & t_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & t_{pp} \end{pmatrix};$$

that $\underset{\sim\sim}{TF'}$ has 0's below the main diagonal follows from the facts that $\underset{\sim}{T}$ and $\underset{\sim}{F'}$ have 0's below the main diagonals, and that $\underset{\sim\sim}{TF'}$ has 0's above the main diagonal follows from the fact that $\underset{\sim\sim}{TF'} = \underset{\sim\sim\sim}{FAF'}$ is symmetric. Since $\underset{\sim\sim\sim}{FAF'}$ is positive definite, $t_{ii} > 0$, $i=1, \ldots, p$. Let

$$(2.15) \qquad \underset{\sim}{D} = \begin{pmatrix} \sqrt{t_{11}} & 0 & 0 & \cdots & 0 \\ 0 & \sqrt{t_{22}} & 0 & \cdots & 0 \\ 0 & 0 & \sqrt{t_{33}} & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \sqrt{t_{pp}} \end{pmatrix},$$

and let $\underset{\sim}{H} = \underset{\sim}{D}^{-1}\underset{\sim}{F}$. Then

$$(2.16) \qquad \underset{\sim}{H}\underset{\sim}{A}\underset{\sim}{H}' = \underset{\sim}{I} ,$$

$$(2.17) \qquad \underset{\sim}{A}^{-1} = \underset{\sim}{H}'\underset{\sim}{H} .$$

(This development is given in more detail by T. W. Anderson (1971) in Section 2.3.) Note that only the forward solution is needed to obtain $\underset{\sim}{H}$.

Then

$$(2.18) \qquad \begin{aligned} \operatorname{tr} \underset{\sim}{A}^{-1} \underset{\sim}{B} \underset{\sim}{A}^{-1} \underset{\sim}{L} &= \operatorname{tr} \underset{\sim}{H}'\underset{\sim}{H} \underset{\sim}{B} \underset{\sim}{H}'\underset{\sim}{H} \underset{\sim}{L} \\ &= \operatorname{tr} \underset{\sim}{H} \underset{\sim}{B} \underset{\sim}{H}'\underset{\sim}{H} \underset{\sim}{L} \underset{\sim}{H}' . \end{aligned}$$

Thus the symmetric matrices $\underset{\sim}{H}\underset{\sim}{B}\underset{\sim}{H}'$ and $\underset{\sim}{H}\underset{\sim}{L}\underset{\sim}{H}'$ are computed. When $N = 1$ and $\underset{\sim}{\mu} = 0$, $\underset{\sim}{C} = \underset{\sim}{x}\underset{\sim}{x}'$ and (2.18) with $\underset{\sim}{L}$ replaced by $\underset{\sim}{C} = \underset{\sim}{x}\underset{\sim}{x}'$ becomes

$$(2.19) \qquad \begin{aligned} \operatorname{tr} \underset{\sim}{H}'\underset{\sim}{H} \underset{\sim}{B} \underset{\sim}{H}'\underset{\sim}{H} \underset{\sim}{x}\underset{\sim}{x}' &= \underset{\sim}{x}' \underset{\sim}{H}'\underset{\sim}{H} \underset{\sim}{B} \underset{\sim}{H}'\underset{\sim}{H} \underset{\sim}{x} \\ &= (\underset{\sim}{H}\underset{\sim}{x})' \underset{\sim}{H} \underset{\sim}{B} \underset{\sim}{H}' (\underset{\sim}{H}\underset{\sim}{x}) \\ &= (\underset{\sim}{F}\underset{\sim}{x})' \underset{\sim}{D}^{-2} \underset{\sim}{F} \underset{\sim}{B} \underset{\sim}{F}' \underset{\sim}{D}^{-2} (\underset{\sim}{F}\underset{\sim}{x}) \\ &= (\underset{\sim}{F}' \underset{\sim}{D}^{-2} \underset{\sim}{F}\underset{\sim}{x}) \underset{\sim}{B} (\underset{\sim}{F}' \underset{\sim}{D}^{-2} \underset{\sim}{F}\underset{\sim}{x}) . \end{aligned}$$

Given $\underset{\sim}{A} = \hat{\underset{\sim}{\Sigma}}_{i-1}$, the data enter the equations (2.2) through $\underset{\sim}{H}\underset{\sim}{C}\underset{\sim}{H}'$. In the case of $\underset{\sim}{C} = \underset{\sim}{x}\underset{\sim}{x}'$, this involves only

$$(2.20) \qquad \underset{\sim\sim}{Hx} = \underset{\sim}{D}^{-1} \underset{\sim}{F} \underset{\sim}{x} ,$$

which is the forward solution applied to $\underset{\sim}{x}$, followed by the division of each element of the resulting vector by the square root of the diagonal element of $\underset{\sim\sim}{FA} = \underset{\sim}{T}$.

If (1.2) holds and $\beta_1, \ldots , \beta_r$ are unknown, initial unbiased estimates of $\beta_1, \ldots , \beta_r$ can be obtained from the equations

$$(2.21) \qquad \sum_{i=1}^{r} \underset{\sim}{z}_j' \underset{\sim}{\Theta} \underset{\sim}{z}_i \hat{\beta}_i = \underset{\sim}{z}_j' \underset{\sim}{\Theta} \underset{\sim}{\bar{x}} , \qquad j=1, \ldots , r ,$$

where $\underset{\sim}{\Theta}$ is any positive definite matrix. The solution has a multivariate normal distribution with covariance matrix

$$(2.22) \qquad \frac{1}{N} (\underset{\sim}{z}_j' \underset{\sim}{\Theta} \underset{\sim}{z}_i)^{-1} (\underset{\sim}{z}_j' \underset{\sim}{\Theta} \underset{\sim}{\Sigma} \underset{\sim}{\Theta} \underset{\sim}{z}_i)(\underset{\sim}{z}_j' \underset{\sim}{\Theta} \underset{\sim}{z}_i)^{-1} .$$

These estimates are consistent as $N \to \infty$. If $\underset{\sim}{\Theta} = \underset{\sim}{I}$, these estimates are least squares.

When $\beta_1, \ldots , \beta_r$ and $\sigma_0, \sigma_1, \ldots , \sigma_m$ are unknown, the estimates obtained from (2.21) can be inserted into (1.5) to obtain an estimate of $\underset{\sim}{\mu}$ and this estimate in turn can be used to define $\hat{\underset{\sim}{C}}$. Then $\sigma_0, \sigma_1, \ldots , \sigma_m$ can be estimated and $\underset{\sim}{\Sigma}$; this estimate of $\underset{\sim}{\Sigma}$ can replace $\underset{\sim}{\Sigma}$ in (1.4) to obtain improved estimates of $\beta_1, \ldots , \beta_r$. This procedure yields consistent and asymptotically efficient estimates of $\beta_1, \ldots , \beta_r, \sigma_0, \sigma_1, \ldots , \sigma_m$ as $N \to \infty$. We shall study later their asymptotic properties as $p \to \infty$.

3. Computation for the Moving Average Process

3.1. Case of First-Order Moving Average Process. The special form of $\underset{\sim}{\Sigma}$ as given in (1.26) makes the computation easier in the case of estimating the nonzero covariances of a finite moving average process. This is of particular interest in the case of one observation $\underset{\sim}{x}$ when $\underset{\sim}{\mu}$ is assumed $\underset{\sim}{0}$. To illustrate the procedure we consider the case of order $m = 1$. Let $\rho = \sigma_1/\sigma_0$, and let the $p \times p$ matrix $(1/\sigma_0)\underset{\sim}{\Sigma}$ be

$$\underset{\sim}{A} = \begin{pmatrix} 1 & \rho & 0 & \cdots & 0 \\ \rho & 1 & \rho & \cdots & 0 \\ 0 & \rho & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} . \tag{3.1}$$

All of the elements of this matrix are 0's except on the main diagonal and one above and below the main diagonal. The method of pivotal condensation or successive elimination starts with leaving the first row unchanged and subtracting ρ times the first row from the second row; this operation changes the second row to a row having nonzero elements only on the diagonal and one entry to the right of the diagonal. Each successive step consists of subtracting a suitable multiple of one row from the next. This "forward solution" can be represented as $\underset{\sim}{F}\underset{\sim}{A}$, where

$$\underset{\sim}{F} = \underset{\sim}{F}_p \underset{\sim}{F}_{p-1} \cdots \underset{\sim}{F}_3 \underset{\sim}{F}_2 \underset{\sim}{F}_1 . \tag{3.2}$$

$$\underset{\sim}{F}_1 = \underset{\sim}{I} . \tag{3.3}$$

(3.4) $$\underset{\sim}{F}_2 = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ -\rho & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix},$$

(3.5) $$\underset{\sim}{F}_3 = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & -\frac{\rho}{1-\rho^2} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}.$$

The matrix $\underset{\sim}{F}_j$ is the identity except that the element in the j-th row and (j-1)-st column is $-\rho$ times the reciprocal of the element in the (j-1)-st row and (j-1)-st column of $\underset{\sim}{F}_{j-1} \underset{\sim}{F}_{j-2} \cdots \underset{\sim}{F}_1 \underset{\sim}{A}$. Thus $\underset{\sim}{F}_j$ has the form

(3.6) $$\underset{\sim}{F}_j = \begin{pmatrix} \underset{\sim}{I} & \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} \\ \underset{\sim}{0} & 1 & \underset{\sim}{0} & \underset{\sim}{0} \\ \underset{\sim}{0} & f_{j,j-1} & 1 & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{I} \end{pmatrix},$$

where the orders of the $\underset{\sim}{I}$'s are $j-2$ and $p-j$, respectively. (In the product $\underset{\sim}{F} = \underset{\sim}{F}_p \underset{\sim}{F}_{p-1} \cdots \underset{\sim}{F}_2 \underset{\sim}{F}_1$ the element $f_{j,j-1}$ appears in the j-th row and (j-1)-st column.) The calculation of $\underset{\sim}{H}\underset{\sim}{x}$ involves

(3.7) $$\underset{\sim\sim}{Fx} = \underset{\sim}{F}_p\, \underset{\sim}{F}_{p-1} \cdots \underset{\sim}{F}_3\, \underset{\sim}{F}_2\, \underset{\sim}{F}_1\, \underset{\sim}{x} = \underset{\sim}{w} ,$$

say. The computation of w proceeds as follows: $w_1 = x_1$,

(3.8) $$w_j = x_j + f_{j,j-1}\, w_{j-1} , \qquad j=2, \ldots , p .$$

Thus the elements of $\underset{\sim}{w}$ can be calculated in sequence.

We can write an equation for $f_{j,j-1}$. Let $a_{jj}^{(j)}$ be the j,j-th term of $\underset{\sim}{F}_j\, \underset{\sim}{F}_{j-1} \cdots \underset{\sim}{F}_1\, \underset{\sim}{A}$. The method of successive elmination shows that

(3.9) $$f_{j+1,j} = -\frac{\rho}{a_{jj}^{(j)}} ,$$

(3.10) $$a_{j+1,j+1}^{(j+1)} = 1 - \frac{\rho^2}{a_{jj}^{(j)}} = \frac{a_{jj}^{(j)} - \rho^2}{a_{jj}^{(j)}} .$$

The elements $a_{j+1,j+1}^{(j+1)}$ and $f_{j+1,j}$, $j=1, 2, \ldots , p-1$, can be computed in sequence. Next let

(3.11) $$\underset{\sim}{D}^{-2}\, \underset{\sim}{w} = \underset{\sim}{u} ,$$

where $\underset{\sim}{D}$ is defined in (2.15); in components this is

(3.12) $$u_j = w_j/d_{jj}^2 = w_j/t_{jj} = w_j\, /a_{jj}^{(j)} , \qquad j=1, \ldots , p .$$

Finally, let

(3.13) $$\underset{\sim}{v} = \underset{\sim}{F}'\, \underset{\sim}{u} = \underset{\sim}{F}_1'\, \underset{\sim}{F}_2' \cdots \underset{\sim}{F}_{p-1}'\, \underset{\sim}{F}_p'\, \underset{\sim}{w} ;$$

thus $v_p = u_p$

(3.14) $$v_j = u_j + f_{j+1,j}\, v_{j+1} , \quad j = p-1, p-2, \ldots , 2, 1 .$$

Then

$$(3.15) \qquad \underset{\sim}{x}' \underset{\sim}{\Sigma}^{-1} \underset{\sim}{G}_h \underset{\sim}{\Sigma}^{-1} \underset{\sim}{x} = \frac{1}{\sigma_0^2} \underset{\sim}{x}' \underset{\sim}{A}^{-1} \underset{\sim}{G}_h \underset{\sim}{A}^{-1} \underset{\sim}{x}$$

$$= \frac{1}{\sigma_0^2} \underset{\sim}{v}' \underset{\sim}{G}_h \underset{\sim}{v} . \qquad h=0,\ 1 .$$

For $\underset{\sim}{G}_0 = \underset{\sim}{I}$, $\underset{\sim}{v}\, \underset{\sim}{G}_0\, \underset{\sim}{v} = \underset{\sim}{v}'\underset{\sim}{v} = \sum_{j=1}^{p} v_j^2$, and for $\underset{\sim}{G}_1$ given by (1.27)

$$(3.16) \qquad \underset{\sim}{v}' \underset{\sim}{G}_1 \underset{\sim}{v} = 2 \sum_{j=1}^{p-1} v_j\, v_{j+1} .$$

The number of arithmetic operations in the calculation of these quadratic forms is approximately proportional to p .

To calculate the coefficients of the unknowns in the iterative procedure we need

$$(3.17) \qquad \operatorname{tr} \underset{\sim}{A}^{-1} \underset{\sim}{G}_g \underset{\sim}{A}^{-1} \underset{\sim}{G}_h = \operatorname{tr} \underset{\sim}{H}'\underset{\sim}{H}\, \underset{\sim}{G}_g\, \underset{\sim}{H}'\underset{\sim}{H}\, \underset{\sim}{G}_h$$

$$= \operatorname{tr} \underset{\sim}{F}' \underset{\sim}{D}^{-2} \underset{\sim}{F} \underset{\sim}{G}_g \underset{\sim}{F}' \underset{\sim}{D}^{-2} \underset{\sim}{F} \underset{\sim}{G}_h$$

$$= \operatorname{tr} (\underset{\sim}{D}^{-1} \underset{\sim}{F} \underset{\sim}{G}_g \underset{\sim}{F}' \underset{\sim}{D}^{-1})(\underset{\sim}{D}^{-1} \underset{\sim}{F} \underset{\sim}{G}_h \underset{\sim}{F}' \underset{\sim}{D}^{-1}) .$$

The forward procedure $\underset{\sim}{F}$ can be applied to each $\underset{\sim}{G}_g$ on the left and $\underset{\sim}{F}'$ on the right followed by $\underset{\sim}{D}^{-1}$ on the right and left to obtain the symmetric matrix $\underset{\sim}{D}^{-1} \underset{\sim}{F} \underset{\sim}{G}_g \underset{\sim}{F}' \underset{\sim}{D}^{-1}$. However, in this case of $m = 1$, the form of $\underset{\sim}{\Sigma}^{-1} = (1/\sigma_0)\underset{\sim}{A}^{-1}$ is known and

$$(3.18) \qquad \operatorname{tr} \underset{\sim}{\Sigma}^{-1} \underset{\sim}{G}_g \underset{\sim}{\Sigma}^{-1} \underset{\sim}{G}_h = \frac{1}{\sigma_0^2} \operatorname{tr} \underset{\sim}{A}^{-1} \underset{\sim}{G}_g \underset{\sim}{A}^{-1} \underset{\sim}{G}_h$$

can be computed directly. See Shaman (1969).

The matrix $\underset{\sim}{G}_1$ can be written

$$(3.19) \qquad \underset{\sim}{G}_1 = \underset{\sim}{P}\, \underset{\sim}{\Lambda}_1\, \underset{\sim}{P}' ,$$

where the orthogonal matrix $\underset{\sim}{P}$ is

$$(3.20)\qquad \underset{\sim}{P} = \sqrt{\frac{2}{p+1}} \left(\sin \frac{\pi jk}{p+1}\right),$$

and the diagonal matrix $\underset{\sim}{\Lambda}$ has $2 \cos \pi j/(p+1)$ as its j-th diagonal element. [See Section 6.5.4 of T. W. Anderson (1971).] Then the quantities (3.18) can also be written

$$(3.21)\qquad \frac{1}{\sigma_0^2} \operatorname{tr} \underset{\sim}{A}^{-1} \underset{\sim}{G}_g \underset{\sim}{A}^{-1} \underset{\sim}{G}_h = \frac{1}{\sigma_0^2} \operatorname{tr}[\underset{\sim}{I} + \rho \underset{\sim}{G}_1]^{-1} \underset{\sim}{G}_g [\underset{\sim}{I} + \rho \underset{\sim}{G}_1]^{-1} \underset{\sim}{G}_h$$

$$= \frac{1}{\sigma_0^2} \operatorname{tr}[\underset{\sim}{P}\underset{\sim}{P}' + \rho \underset{\sim}{P} \underset{\sim}{\Lambda}_1 \underset{\sim}{P}']^{-1} \underset{\sim}{P} \underset{\sim}{\Lambda}_g \underset{\sim}{P}' [\underset{\sim}{P}\underset{\sim}{P}' + \rho \underset{\sim}{P} \underset{\sim}{\Lambda}_1 \underset{\sim}{P}']^{-1} \underset{\sim}{P} \underset{\sim}{\Lambda}_h \underset{\sim}{P}'$$

$$= \frac{1}{\sigma_0^2} \operatorname{tr}[\underset{\sim}{P}(\underset{\sim}{I} + \rho \underset{\sim}{\Lambda}_1)\underset{\sim}{P}']^{-1} \underset{\sim}{P} \underset{\sim}{\Lambda}_g \underset{\sim}{P}' [\underset{\sim}{P}(\underset{\sim}{I} + \rho \underset{\sim}{\Lambda}_1)\underset{\sim}{P}']^{-1} \underset{\sim}{P} \underset{\sim}{\Lambda}_h \underset{\sim}{P}'$$

$$= \frac{1}{\sigma_0^2} \operatorname{tr} \underset{\sim}{P}(\underset{\sim}{I} + \rho \underset{\sim}{\Lambda}_1)^{-1} \underset{\sim}{P}'\underset{\sim}{P} \underset{\sim}{\Lambda}_g \underset{\sim}{P}'\underset{\sim}{P}(\underset{\sim}{I} + \rho \underset{\sim}{\Lambda}_1)^{-1} \underset{\sim}{P}'\underset{\sim}{P} \underset{\sim}{\Lambda}_h \underset{\sim}{P}$$

$$= \frac{1}{\sigma_0^2} \operatorname{tr}(\underset{\sim}{I} + \rho \underset{\sim}{\Lambda}_1)^{-1} \underset{\sim}{\Lambda}_g (\underset{\sim}{I} + \rho \underset{\sim}{\Lambda}_1)^{-1} \underset{\sim}{\Lambda}_h, \qquad g, h=0, 1,$$

where $\underset{\sim}{\Lambda}_0 = \underset{\sim}{I}$. Then

$$(3.22)\qquad \operatorname{tr} \underset{\sim}{A}^{-1} \underset{\sim}{G}_0 \underset{\sim}{A}^{-1} \underset{\sim}{G}_0 = \sum_{j=1}^{p} \frac{1}{(1+2\rho \cos \frac{\pi j}{p+1})^2},$$

$$(3.23)\qquad \operatorname{tr} \underset{\sim}{A}^{-1} \underset{\sim}{G}_0 \underset{\sim}{A}^{-1} \underset{\sim}{G}_1 = \sum_{j=1}^{p} \frac{2 \cos \frac{\pi j}{p+1}}{(1+2\rho \cos \frac{\pi j}{p+1})^2},$$

$$(3.24)\qquad \operatorname{tr} \underset{\sim}{A}^{-1} \underset{\sim}{G}_1 \underset{\sim}{A}^{-1} \underset{\sim}{G}_1 = \sum_{j=1}^{p} \frac{4 \cos^2 \frac{\pi j}{p+1}}{(1+2\rho \cos \frac{\pi j}{p+1})^2},$$

These sums can be approximated by integrals. The sum (3.22) is approximated by

(3.25) $$\frac{p+1}{\pi} \int_{\frac{\pi}{2(p+1)}}^{\pi - \frac{\pi}{2(p+1)}} \frac{d\lambda}{(1+2\rho \cos \lambda)^2} ,$$

which in turn is approximated by

(3.26) $$\frac{p+1}{\pi} \left\{ \int_0^{\pi} \frac{d\lambda}{(1+2\rho \cos \lambda)^2} - \frac{\pi}{2(p+1)} \left[\frac{1}{(1+2\rho)^2} + \frac{1}{(1-2\rho)^2} \right] \right\}$$

$$= \frac{p+1}{\pi} \int_0^{\pi} \frac{d\lambda}{(1+2\rho \cos \lambda)^2} - \frac{1+4\rho^2}{(1-4\rho^2)^2} .$$

The first term on the right-hand side of (3.26) is $(p+1)/(1-4\rho^2)^{3/2}$. [See Pierce (1929), Formulas 300 and 305, for example.] Then tr $\underset{\sim}{\Sigma}^{-2}$ is approximated by

(3.27) $$\frac{1}{\sigma^4} \left[\frac{(p+1)(1+\alpha^2)}{(1-\alpha^2)^3} - \frac{1+6\alpha^2 + \alpha^4}{(1-\alpha^2)^4} \right]$$

$$= \frac{1}{\sigma^4} \; \frac{p - 6\alpha^2 - (p+2)\alpha^4}{(1-\alpha^2)^4} .$$

In a similar way (3.23) and (3.24) are approximated by

(3.28) $$- \frac{4(p+1)\rho}{(1-4\rho^2)^{3/2}} + \frac{8\rho}{(1-4\rho^2)^2} ,$$

(3.29) $$\frac{4(p+1)}{\rho^2} \left[1 - \frac{1-8\rho^2}{(1-4\rho^2)^{3/2}} \right] - 4 \, \frac{1+4\rho^2}{(1-4\rho^2)^2} , \qquad \rho \neq 0 ,$$

$$4p , \qquad \rho = 0 .$$

The 2 x 2 matrix of coefficients $\operatorname{tr} \underset{\sim}{\Sigma}^{-1} \underset{\sim}{G}_g \underset{\sim}{\Sigma}^{-1} \underset{\sim}{G}_h$ (g, h=0, 1) is approximated by

$$(3.30) \qquad \frac{p+1}{\sigma^4 (1-\alpha^2)^3} \begin{pmatrix} 1+\alpha^2 & -4\alpha \\ -4\alpha & 2+8\alpha^2 - 2\alpha^4 \end{pmatrix} - \frac{1}{\sigma^4 (1-\alpha^2)^4} \begin{pmatrix} 1+6\alpha^2 + \alpha^4 & -8\alpha(1+\alpha^2) \\ -8\alpha(1+\alpha^2) & 4(1+\alpha^2) \end{pmatrix} .$$

An alternative approach, which may be generalized to cases of $m > 1$, is to approximate the moving average covariance matrix by the inverse of an autoregressive covariance matrix. The covariance matrix of the moving average process of order 1 may be written

$$(3.31) \qquad \underset{\sim}{\Sigma}_{MA} = \sigma^2 \begin{pmatrix} 1+\alpha^2 & \alpha & 0 & \cdots & 0 & 0 \\ \alpha & 1+\alpha^2 & \alpha & \cdots & 0 & 0 \\ 0 & \alpha & 1+\alpha^2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1+\alpha^2 & \alpha \\ 0 & 0 & 0 & \cdots & \alpha & 1+\alpha^2 \end{pmatrix} .$$

The matrix in the exponent of the first-order autoregressive Gaussian process satisfying the stochastic difference equation

$$(3.32) \qquad y_t + \alpha y_{t-1} = u_t , \qquad t = \ldots -1, 0, 1, \ldots ,$$

is

$$(3.33)\qquad \underset{\sim}{\Psi}_{AR} = \frac{1}{\sigma_u^2}\begin{pmatrix} 1 & \alpha & 0 & \cdots & 0 & 0 \\ \alpha & 1+\alpha^2 & \alpha & \cdots & 0 & 0 \\ 0 & \alpha & 1+\alpha^2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1+\alpha^2 & \alpha \\ 0 & 0 & 0 & \cdots & \alpha & 1 \end{pmatrix} .$$

This differs from $\underset{\sim}{\Sigma}_{MA}$ by $\sigma^2\alpha^2\ \underset{\sim}{E}$, where

$$(3.34)\qquad \underset{\sim}{E} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix} ,$$

if $\sigma^2 = 1/\sigma_u^2$. This fact suggests approximating $\underset{\sim}{\Sigma}_{MA}^{-1}$ by

$$(3.35)\qquad \underset{\sim}{\Sigma}_{AR} = \underset{\sim}{\Psi}_{AR}^{-1} = \frac{\sigma_u^2}{1-\alpha^2}\begin{pmatrix} 1 & -\alpha & \alpha^2 & \cdots & (-\alpha)^{p-2} & (-\alpha)^{p-1} \\ -\alpha & 1 & -\alpha & \cdots & (-\alpha)^{p-3} & (-\alpha)^{p-2} \\ \alpha^2 & -\alpha & 1 & \cdots & (-\alpha)^{p-4} & (-\alpha)^{p-3} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ (-\alpha)^{p-2} & (-\alpha)^{p-3} & (-\alpha)^{p-4} & \cdots & 1 & -\alpha \\ (-\alpha)^{p-1} & (-\alpha)^{p-2} & (-\alpha)^{p-3} & \cdots & -\alpha & 1 \end{pmatrix}$$

with $\sigma_u^2 = 1/\sigma^2$. Then

$$(3.36)\quad \underset{\sim}{\Sigma}_{AR}\, \underset{\sim}{G}_0\, \underset{\sim}{\Sigma}_{AR} = \frac{1}{\sigma^4\ (1-\alpha^2)^3}$$

$$\begin{pmatrix} 1+\alpha^2-\alpha^2-\alpha^{2p} & -\alpha[2-\alpha^2-\alpha^{2p-2}] & \alpha^2[3-2\alpha^2-\alpha^{2p-4}] & \cdots & \\ -\alpha[2-\alpha^2-\alpha^{2p-2}] & 1+\alpha^2-\alpha^4-\alpha^{2p-2} & -\alpha[2-\alpha^4-\alpha^{2p-4}] & \cdots & \\ \alpha^2[3-2\alpha^2-\alpha^{2p-4}] & -\alpha[2-\alpha^4-\alpha^{2p-4}] & 1+\alpha^2-\alpha^6-\alpha^{2p-4} & \cdots & \\ \vdots & \vdots & \vdots & & \\ (-\alpha)^{p-1}[p-(p-1)\alpha^2-\alpha^2] & & & & 1+\alpha^2-\alpha^2-\alpha^{2p} \end{pmatrix}$$

and

$$(3.37)\quad \mathrm{tr}\ \underset{\sim}{\Sigma}_{AR}\, \underset{\sim}{G}_0\, \underset{\sim}{\Sigma}_{AR}\, \underset{\sim}{G}_0 = \frac{1}{\sigma^4\ (1-\alpha^2)^4}\ \{p(1-\alpha^4)-2\alpha^2+2\alpha^{2p+2}\}\ .$$

The other quantities needed $\underset{\sim}{\Sigma}_{AR}\, \underset{\sim}{G}_1\, \underset{\sim}{\Sigma}_{AR}$, $\mathrm{tr}\ \underset{\sim}{\Sigma}_{AR}\, \underset{\sim}{G}_1\, \underset{\sim}{\Sigma}_{AR}\, \underset{\sim}{G}_0$, and $\mathrm{tr}\ \underset{\sim}{\Sigma}_{AR}\, \underset{\sim}{G}_1\, \underset{\sim}{\Sigma}_{AR}\, \underset{\sim}{G}_1$ can also be computed.

In the iterative procedure $\hat{\underset{\sim}{\Sigma}}{}_{i-1}^{-1}$ is replaced by $\underset{\sim}{\Sigma}_{AR}$ where $\sigma^2(1+\alpha^2)$ is $\hat{\sigma}_0^{(i-1)}$ and $\sigma^2\alpha$ is $\hat{\sigma}_1^{(i-1)}$. In situations where p is large and we consider limits as $p\to\infty$, we can use

$$(3.38)\qquad \lim_{p\to\infty} \frac{\mathrm{tr}\ \underset{\sim}{\Sigma}_{AR}\, \underset{\sim}{G}_0\, \underset{\sim}{\Sigma}_{AR}\, \underset{\sim}{G}_0}{p} = \frac{1}{\sigma^4\ (1-\alpha^2)^3}\ ,$$

$$(3.39)\qquad \lim_{p\to\infty} \frac{\mathrm{tr}\ \underset{\sim}{\Sigma}_{AR}\, \underset{\sim}{G}_1\, \underset{\sim}{\Sigma}_{AR}\, \underset{\sim}{G}_1}{p} = \frac{2+8\alpha^2-2\alpha^4}{\sigma^4\ (1-\alpha^2)^3}\ ,$$

$$(3.40)\qquad \lim_{p\to\infty} \frac{\mathrm{tr}\ \underset{\sim}{\Sigma}_{AR}\, \underset{\sim}{G}_0\, \underset{\sim}{\Sigma}_{AR}\, \underset{\sim}{G}_1}{p} = \frac{-4\alpha}{\sigma^4\ (1-\alpha^2)^3}\ .$$

If the matrix A given by (3.1) is used in place of $\underset{\sim}{\Sigma} = \Sigma_{\sim AR}$, then σ^2 can be replaced by 1 and $\alpha/(1+\alpha^2)$ by ρ in (3.38), (3.39), and (3.40) to give $1/(1-4\rho^2)^{3/2}$ and the coefficients of $p+1$ in (3.28) and (3.29), respectively. These values agree with the limit of $1/p$ times (3.30).

$$r_1 = \frac{C_1}{C_0}, \tag{3.41}$$

where

$$\begin{aligned} C_h &= \frac{1}{T-h} \sum_{j=1}^{T-h} x_j \, x_{j+h} \\ &= \frac{\underset{\sim}{x}' \, \underset{\sim}{G}_h \, \underset{\sim}{x}}{\operatorname{tr} \underset{\sim}{G}_h^2} \end{aligned} \tag{3.42}$$

This is a consistent estimate of ρ as $p \to \infty$. The estimates C_0 and C_1 are the unbiased estimates of σ_0 and σ_1 obtained from (2.6) with $\underset{\sim}{\Theta} = \underset{\sim}{I}$.

3.2. Case of Higher-Order Moving Average Process. Let $\rho_h = \sigma_h/\sigma_0$ and let the $p \times p$ matrix $(1/\sigma_0)\underset{\sim}{\Sigma}$ be

$$\underset{\sim}{A} = \begin{pmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_m & 0 & \cdots & 0 \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{m-1} & \rho_m & \cdots & 0 \\ \rho_2 & \rho_1 & 1 & \cdots & \rho_{m-2} & \rho_{m-1} & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \rho_m & \rho_{m-1} & \rho_{m-2} & \cdots & 1 & \rho_1 & \cdots & 0 \\ 0 & \rho_m & \rho_{m-1} & \cdots & \rho_1 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 1 \end{pmatrix}, \tag{3.43}$$

which has 0's $m+1$ and more positions above and below the main diagonal ($m < p$). Consider again the forward solution of a method of pivotal condensation or successive elimination. We represent this as the multiplication of A on the left by $\underset{\sim}{F} = \underset{\sim}{F}_p \underset{\sim}{F}_{p-1} \cdots \underset{\sim}{F}_2 \underset{\sim}{F}_1$, where $\underset{\sim}{F}_1 = \underset{\sim}{I}$. In one procedure $\underset{\sim}{F}_2$ represents subtracting ρ_1 times the first row of $\underset{\sim}{F}_1 \underset{\sim}{A} = \underset{\sim}{A}$ from the second, ρ_2 times the first row from the third, ..., and ρ_m times the first row from the $(m+1)$st. Then $\underset{\sim}{F}_2 \underset{\sim}{F}_1 \underset{\sim}{A}$ has all 0's in the first column below the first entry. Before the j-th step $\underset{\sim}{F}_{j-1} \cdots \underset{\sim}{F}_1 \underset{\sim}{A}$ has all 0's in the first $j-1$ columns below the main diagonal. $\underset{\sim}{F}_j$ represents subtracting multiples of the $(j-1)$st row of $\underset{\sim}{F}_{j-1} \cdots \underset{\sim}{F}_1 \underset{\sim}{A}$ from the j-th, $(j+1)$st, ..., $(j+m-1)$st rows in order that the j-th column have only 0's below the main diagonal. Let $a_{ik}^{(j-1)}$ be the i, k-th element of $\underset{\sim}{A}^{(j-1)} = \underset{\sim}{F}_{j-1} \cdots \underset{\sim}{F}_1 \underset{\sim}{A}$. Then

$$(3.44) \qquad \underset{\sim}{F}_j = \begin{pmatrix} \underset{\sim}{I} & \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} \\ \underset{\sim}{0} & 1 & \underset{\sim}{0} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{f}_j & \underset{\sim}{I} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{I} \end{pmatrix},$$

where the $\underset{\sim}{I}$'s are of order $j-2$, m, and $p+1-j-m$, respectively, and

$$(3.45) \qquad \underset{\sim}{f}_j = -\begin{pmatrix} -\dfrac{\rho_1}{a_{j-1,j-1}^{(j-1)}} \\ \vdots \\ -\dfrac{\rho_m}{a_{j-1,j-1}^{(j-1)}} \end{pmatrix}, \qquad j=2, \ldots, p-m+1 ;$$

for $j = p-m+2, \ldots, p$, $\underset{\sim}{f}_j$ consists of the first $p-j+1$ components of (3.45). Thus $\underset{\sim}{F}_j$ consists of the identity with (at most) m nonzero elements below the main diagonal in the $(j-1)$st column.

The operations can be done in another order. Then $\underset{\sim}{F}_2$ represents subtracting ρ_1 times the first row of $\underset{\sim}{F}_1 \underset{\sim}{A} = \underset{\sim}{A}$ from the second; $\underset{\sim}{F}_3$ represents subtracting appropriate multiples of the first two rows of $\underset{\sim}{F}_2 \underset{\sim}{F}_1 \underset{\sim}{A}$ from the third row to obtain 0's to the left of the diagonal elements in the third row of $\underset{\sim}{F}_3 \underset{\sim}{F}_2 \underset{\sim}{F}_1 \underset{\sim}{A}$. Before the j-th step $\underset{\sim}{F}_{j-1} \cdots \underset{\sim}{F}_1 \underset{\sim}{A}$ has 0's in the first $j-1$ rows to the left of the main diagonal and the remaining $p-j+1$ rows of $\underset{\sim}{F}_{j-1} \cdots \underset{\sim}{F}_1 \underset{\sim}{A}$ are the same as those of $\underset{\sim}{A}$. $\underset{\sim}{F}_j$ represents subtracting appropriate multiples of the $(j-m)$th, $(j-m+1)$st, ..., $(j-1)$st rows of $\underset{\sim}{F}_{j-1} \cdots \underset{\sim}{F}_1 \underset{\sim}{A}$ in turn to make the $(j-m)$th, $(j-m+1)$st, ..., $(j-1)$st elements of the j-th row of $\underset{\sim}{F}_j \cdots \underset{\sim}{F}_1 \underset{\sim}{A}$ zero. Thus $\underset{\sim}{F}_j$ consists of the identity with (at most) m nonzero elements to the left of the main diagonal in the j-th row.

The calculation of $\underset{\sim}{H}\underset{\sim}{x}$ is done by

$$\underset{\sim}{w} = \underset{\sim}{F}\underset{\sim}{x} = \underset{\sim}{F}_p \underset{\sim}{F}_{p-1} \cdots \underset{\sim}{F}_2 \underset{\sim}{F}_1 \underset{\sim}{x} . \tag{3.46}$$

Let

$$\underset{\sim}{w}^{(j)} = \underset{\sim}{F}_j \underset{\sim}{w}^{(j-1)} . \tag{3.47}$$

Then $\underset{\sim}{w} = \underset{\sim}{w}^{(p)}$. In the first sequence of operations in (3.47) $w_i^{(j)} = w_i^{(j-1)}$ for $i=1, \ldots, j-1$,

$$(3.48) \qquad w_i^{(j)} = w_i^{(j-1)} - f_{i,j-1}^{(j)} \, w_{j-1}^{(j-1)} , \qquad i=j, \ldots , j+m-1 ,$$

$w_i^{(j)} = w_i^{(j-1)} = x_i$, $i = j+m, \ldots , p$; if $j = p-m+2, \ldots , p$ (3.48) holds for $i=j, \ldots , p$. Thus $\underset{\sim}{w}^{(j)}$ [illegible]omputed successively.

Next let

$$(3.49) \qquad \underset{\sim}{u} = \underset{\sim}{D}^{-2} \underset{\sim}{w}$$

and

$$(3.50) \qquad \underset{\sim}{v} = \underset{\sim}{F}' \underset{\sim}{u} = \underset{\sim}{F}_1' \underset{\sim}{F}_2' \cdots \underset{\sim}{F}_{p-1}' \underset{\sim}{F}_p' \underset{\sim}{u} .$$

The operation (3.50) can be done sequentially

$$(3.51) \qquad \underset{\sim}{v}^{(j)} = \underset{\sim}{F}_j' \underset{\sim}{v}^{(j+1)} , \qquad j=p, \ldots , 1 ,$$

where $\underset{\sim}{v}^{(p+1)} = \underset{\sim}{u}$ and $\underset{\sim}{v}^{(1)} = \underset{\sim}{v}$. Then

$$(3.52) \qquad \underset{\sim}{x}' \underset{\sim}{\Sigma}^{-1} \underset{\sim}{G}_h \underset{\sim}{\Sigma}^{-1} \underset{\sim}{x} = \frac{1}{\sigma_0^2} \underset{\sim}{x}' \underset{\sim}{A}^{-1} \underset{\sim}{G}_h \underset{\sim}{A}^{-1} \underset{\sim}{x}$$

$$= \frac{1}{\sigma_0^2} \underset{\sim}{v}' \underset{\sim}{G}_h \underset{\sim}{v} , \qquad h=0, 1, \ldots , m .$$

For $\underset{\sim}{G}_0 = \underset{\sim}{I}$, $\underset{\sim}{v}' \underset{\sim}{G}_0 v = \sum_{j=1}^{p} \underset{\sim}{v}_j^2$ and for $\underset{\sim}{G}_h$ given by (1.27), (1.28), etc.,

$$(3.53) \qquad \underset{\sim}{v}' \underset{\sim}{G}_h \underset{\sim}{v} = 2 \sum_{i=1}^{p-h} v_j \, v_{j+h} , \qquad h=1, \ldots , m .$$

The number of arithmetic operations is approximately proportional to mp .

As indicated for the case $m=1$ the coefficients of $\hat{\sigma}_0, \hat{\sigma}_1, \ldots, \hat{\sigma}_m$ can be calculated from the forward solution according to (3.17) and (3.18).

The approximations to $\operatorname{tr} \underset{\sim}{A}^{-1} \underset{\sim}{G}_g \underset{\sim}{A}^{-1} \underset{\sim}{G}_h$ presented for the case $m=1$ can be extended. The matrix $\underset{\sim}{G}_h$ is approximately $2 \underset{\sim}{A}_h$ treated in Section 6.5.4 of T. W. Anderson (1971). Then $\underset{\sim}{G}_h$ is approximately $\underset{\sim}{P} \underset{\sim}{\Lambda}_h \underset{\sim}{P}$, where $\underset{\sim}{\Lambda}_h$ is a diagonal matrix with $2 \cos hj\pi/(p+1)$ as its j-th diagonal elements. ($\underset{\sim}{P}' \underset{\sim}{G}_0 \underset{\sim}{P}' = \underset{\sim}{P}'\underset{\sim}{P} = \underset{\sim}{I}$, $\underset{\sim}{P}' \underset{\sim}{G}_1 \underset{\sim}{P}' = \underset{\sim}{\Lambda}_1$ diagonal and $\underset{\sim}{P}' \underset{\sim}{G}_h \underset{\sim}{P}$ is $\underset{\sim}{\Lambda}_h$ diagonal plus a matrix relatively small, $j=2, \ldots, m$.) Then we have the approximations

$$(3.54) \quad \operatorname{tr} \underset{\sim}{A}^{-1} \underset{\sim}{G}_0 \underset{\sim}{A}^{-1} \underset{\sim}{G}_0 \sim \sum_{j=1}^{p} \frac{1}{\left(1 + 2 \sum_{h=1}^{m} \rho_h \cos \frac{hj\pi}{p+1}\right)^2}$$

$$\sim \frac{p+1}{\pi} \int_{\frac{\pi}{2(p+1)}}^{\pi - \frac{\pi}{2(p+1)}} \frac{d\lambda}{\left(1 + 2 \sum_{h=1}^{m} \rho_h \cos \lambda h\right)^2},$$

$$(3.55) \quad \operatorname{tr} \underset{\sim}{A}^{-1} \underset{\sim}{G}_0 \underset{\sim}{A}^{-1} \underset{\sim}{G}_g \sim 2 \sum_{j=1}^{p} \frac{\cos \frac{gj\pi}{T+1}}{\left(1 + 2 \sum_{h=1}^{m} \cos \frac{hj\pi}{p+1}\right)^2}$$

$$\sim 2 \frac{p+1}{\pi} \int_{\frac{\pi}{2(p+1)}}^{\pi - \frac{\pi}{2(p+1)}} \frac{\cos \lambda g \, d\lambda}{\left(1 + 2 \sum_{h=1}^{m} \cos \lambda h\right)^2},$$

$$g=1, \ldots, m,$$

$$(3.56)\qquad \operatorname{tr} \underset{\sim}{A}^{-1} \underset{\sim}{G}_f \underset{\sim}{A}^{-1} \underset{\sim}{G}_g \sim 4 \sum_{j=1}^{p} \frac{\cos \frac{fj\pi}{T+1} \cos \frac{gj\pi}{T+1}}{\left(1 + 2 \sum_{h=1}^{m} \cos \frac{hj\pi}{p+1}\right)^2}$$

$$\sim 4 \frac{p+1}{\pi} \int_{\frac{\pi}{2(p+1)}}^{\pi - \frac{\pi}{2(p+1)}} \frac{\cos \lambda f \cos \lambda g \, d\lambda}{\left(1 + 2 \sum_{h=1}^{m} \cos \lambda h\right)^2},$$

$$f, g = 1, \ldots, m .$$

In the integrals $\cos k$ can be written as a polynomial in $\cos \lambda$ of degree k.

As in the case of $m=1$, the covariance matrix $\underset{\sim}{\Sigma}_{MA}$ of the moving average process of order m can be approximated by $\underset{\sim}{\Psi}_{AR}$, the matrix in the exponent of the normal distribution of the Gaussian autoregressive process of order m. Then $\underset{\sim}{\Sigma}_{MA}^{-1}$ is approximated by $\underset{\sim}{\Psi}_{AR}^{-1} = \underset{\sim}{\Sigma}_{AR}$, whose elements are the covariances of the autoregressive process. If the roots of (1.32) are different, say, $z_1, \ldots, z_m$ then the i,j-th element of $\underset{\sim}{\Sigma}_{AR}$ can be written

$$(3.57)\qquad \sigma_{AR}(i-j) = \sum_{h=1}^{m} k_h \, z_h^{|i-j|},$$

for suitable constants $k_1, \ldots, k_h$. [See T. W. Anderson (1971), Section 5.2.2.] Then the i, i-th element of $\underset{\sim}{\Sigma}_{AR}^2$ can be written

$$(3.58)\qquad \sum_{j=1}^{p} \sigma_{AR}^2(i-j) = \sum_{j=1}^{i-1} \sum_{g,h=1}^{m} k_g k_h z_g^{i-j} z_h^{i-j}$$

$$+ \sum_{j=1}^{p} \sum_{g,h=1}^{m} k_g k_h y_g^{j-i} y_h^{j-i}$$

$$= \sum_{g,h=1}^{m} k_g k_h \frac{[1-(z_g z_h)^{i-1}] z_g z_h + 1-(z_g z_h)^{p-i+1}}{1 - z_g z_h}$$

and

(3.59) $$\operatorname{tr} \underset{\sim}{\Sigma}_{AR}^2 = \sum_{i,j=1}^{p} \sigma_{AR}^2(i-j)$$

$$= \sum_{g,h=1}^{m} k_g k_h \frac{p(1 + z_g z_h) - 2 z_g z_h (1 - z_g z_h)/[1 - (z_g z_h)^p]}{1 - z_g z_h}$$

Then

(3.60) $$\lim_{p\to\infty} \frac{1}{p} \operatorname{tr} \underset{\sim}{\Sigma}_{AR}^2 = \sum_{g,h=1}^{p} k_g k_h \frac{1 + z_g z_h}{1 - z_g z_h} .$$

This gives a method for approximating $\operatorname{tr} \underset{\sim}{\Sigma}^{-1} \underset{\sim}{G}_g \underset{\sim}{\Sigma}^{-1} \underset{\sim}{G}_h$. If the integrals in (3.54), (3.55) and (3.56) are over the interval $-\pi, \pi$, they are identical to the evaulation of $\operatorname{tr} \underset{\sim}{\Sigma}_{AR} \underset{\sim}{G}_g \underset{\sim}{\Sigma}_{AR} \underset{\sim}{G}_h$ for the integrand in (3.54) is proportional to the square of the spectral density of the autoregressive process. (See Theorem 8.3.3 of T. W. Anderson (1971).]

We primarily want to use this approximation to obtain the coefficients of $\hat{\sigma}_0^{(i)}, \hat{\sigma}_1^{(i)}, \ldots, \hat{\sigma}_m^{(i)}$ in (2.1) on the basis of $\hat{\sigma}_0^{(i-1)}$, $\hat{\sigma}_1^{(i-1)}, \ldots, \hat{\sigma}_m^{(i-1)}$. That matrix of coefficients is a consistent estimate of 2 times the inverse of the covariance matrix of the asymptotic normal distribution of the asymptotically efficient estimates. [See (1.10).]

<u>Theorem 3.1</u>. <u>If</u> $f(\lambda)$ <u>given by</u> (1.31) <u>is positive, then</u>

(3.58) $$\lim_{p\to\infty} \frac{1}{p} \operatorname{tr} \underset{\sim}{\Sigma}^{-2} = \frac{1}{(2\pi)^2} \int_{-\pi}^{\pi} \frac{1}{f^2(\lambda)} d\lambda ,$$

(3.66) $$\sum_{g,h=0}^{m} \operatorname{tr} \underset{\sim}{A}\, \underset{\sim}{G}_g\, \underset{\sim}{A}\, \underset{\sim}{G}_h\, z_g z_h \le \sum_{g,h=0}^{m} \operatorname{tr} \underset{\sim}{L}\, \underset{\sim}{G}_g\, \underset{\sim}{L}\, \underset{\sim}{G}_h\, z_g z_h$$

for any symmetric matrices $\underset{\sim}{G}_0, \underset{\sim}{G}_1, \ldots, \underset{\sim}{G}_m$.

<u>Proof</u>. Let $\underset{\sim}{F}$ be a nonsingular matrix such that

(3.67) $$\underset{\sim}{L} = \underset{\sim}{F}\underset{\sim}{F}' ,$$

(3.68) $$\underset{\sim}{A} = \underset{\sim}{F}\underset{\sim}{\Delta}\underset{\sim}{F}' ,$$

where $\underset{\sim}{\Delta}$ is diagonal with diagonal elements $\delta_{11}, \ldots, \delta_{pp}$. [See Problem 30, Chapter 6, T. W. Anderson (1971) or Theorem 3, page 341, T. W. Anderson (1958).] Then

(3.69) $$0 \le \delta_{ii} \le 1 , \qquad i=1, \ldots, p .$$

Let $\underset{\sim}{H} = \sum_{g=0}^{m} z_g \underset{\sim}{G}_g$. Then

(3.70) $$\begin{aligned} \sum_{g,h=0}^{m} \operatorname{tr} \underset{\sim}{L}\, \underset{\sim}{G}_g\, \underset{\sim}{L}\, \underset{\sim}{G}_h\, z_g z_h &= \operatorname{tr} \underset{\sim}{F}\underset{\sim}{F}'\, \underset{\sim}{H}\, \underset{\sim}{F}\underset{\sim}{F}'\, \underset{\sim}{H} \\ &= \operatorname{tr} \underset{\sim}{F}'\, \underset{\sim}{H}\, \underset{\sim}{F}\underset{\sim}{F}'\, \underset{\sim}{H}\, \underset{\sim}{F} \\ &= \operatorname{tr} \underset{\sim}{K}^2 \\ &= \sum_{i,j=1}^{p} k_{ij}^2 , \end{aligned}$$

where $\underset{\sim}{K} = \underset{\sim}{F}'\, \underset{\sim}{H}\, \underset{\sim}{F}$. Similarly

(3.71) $$\begin{aligned} \sum_{g,h=0}^{m} \operatorname{tr} \underset{\sim}{A}\, \underset{\sim}{G}_g\, \underset{\sim}{A}\, \underset{\sim}{G}_h\, z_g z_h &= \operatorname{tr} \underset{\sim}{F}\, \underset{\sim}{\Delta}\, \underset{\sim}{F}'\, \underset{\sim}{H}\, \underset{\sim}{F}\, \underset{\sim}{\Delta}\, \underset{\sim}{F}'\, \underset{\sim}{H} \\ &= \operatorname{tr} \underset{\sim}{\Delta}\, \underset{\sim}{F}'\, \underset{\sim}{H}\, \underset{\sim}{F}\, \underset{\sim}{\Delta}\, \underset{\sim}{F}'\, \underset{\sim}{H}\, \underset{\sim}{F} \\ &= \operatorname{tr} \underset{\sim}{\Delta}\, \underset{\sim}{K}\, \underset{\sim}{\Delta}\, \underset{\sim}{K} \\ &= \sum_{i,j=1}^{p} \delta_{ii}\, k_{ij}^2\, \delta_{jj} . \end{aligned}$$

$$(3.59)\qquad \lim_{p\to\infty} \frac{1}{p} \operatorname{tr} \underset{\sim}{\Sigma}^{-1} \underset{\sim}{G}_g \underset{\sim}{\Sigma}^{-1} = \frac{1}{(2\pi)^2} \int_{-\pi}^{\pi} \frac{\cos \lambda g}{f^2(\lambda)}\, d\lambda , \quad g=1, \ldots , m ,$$

$$(3.60)\qquad \lim_{p\to\infty} \frac{1}{p} \operatorname{tr} \underset{\sim}{\Sigma}^{-1} \underset{\sim}{G}_g \underset{\sim}{\Sigma}^{-1} \underset{\sim}{G}_h = \frac{1}{(2\pi)^2} \int_{-\pi}^{\pi} \frac{\cos \lambda g \cos \lambda h}{f^2(\lambda)}\, d\lambda ,$$

$$g, h=1, \ldots , m .$$

Proof. The proof is along the lines of the proof of Theorem 10.2.7 of T. W. Anderson (1971). The spectral density $f(\lambda)$ is continuous. Therefore, for arbitrary $\varepsilon > 0$ there exist autoregressive processes with covariance matrices $\underset{\sim}{\Sigma}_L$ and $\underset{\sim}{\Sigma}_U$ and (positive) spectral densities $f_L(\lambda)$ and $f_U(\lambda)$, respectively, such that

$$(3.61)\qquad f_L(\lambda) \le f(\lambda) \le f_U(\lambda) , \qquad -\pi \le \lambda \le \pi ,$$

$$(3.62)\qquad \frac{1}{f_L(\lambda)} - \frac{1}{f_U(\lambda)} \le \varepsilon \qquad -\pi \le \lambda \le \pi .$$

Then, by Lemma 10.2.6 of T. W. Anderson (1971),

$$(3.63)\qquad \underset{\sim}{x}' \underset{\sim}{\Sigma}_L \underset{\sim}{x} \le \underset{\sim}{x}' \underset{\sim}{\Sigma} \underset{\sim}{x} \le \underset{\sim}{x}' \underset{\sim}{\Sigma}_U \underset{\sim}{x} ,$$

$$(3.64)\qquad \underset{\sim}{x}' \underset{\sim}{\Sigma}_U^{-1} \underset{\sim}{x} \le \underset{\sim}{x}' \underset{\sim}{\Sigma}^{-1} \underset{\sim}{x} \le \underset{\sim}{x}' \underset{\sim}{\Sigma}_L^{-1} \underset{\sim}{x} ,$$

for all $\underset{\sim}{x}$; here $\underset{\sim}{\Sigma}_L$ and $\underset{\sim}{\Sigma}_U$ are the $p \times p$ covariance matrices corresponding to $f_L(\lambda)$ and $f_U(\lambda)$, respectively.

Lemma 3.1. If for $\underset{\sim}{A}$ positive semidefinite and $\underset{\sim}{L}$ positive definite

$$(3.65)\qquad \underset{\sim}{x}' \underset{\sim}{A} \underset{\sim}{x} \le \underset{\sim}{x}' \underset{\sim}{L} \underset{\sim}{x}$$

for all $\underset{\sim}{x}$, then

Since $\delta_{ii} \leq 1$, (3.71) is less than or equal to (3.70). This proves the lemma.

It follows from the lemma that

$$\text{(3.72)} \qquad \sum_{g,h=0}^{m} \operatorname{tr} \underset{\sim}{\Sigma}_U^{-1} \underset{\sim}{G}_g \underset{\sim}{\Sigma}_U^{-1} y_g y_h \leq \sum_{g,h=0}^{m} \operatorname{tr} \underset{\sim}{\Sigma}^{-1} \underset{\sim}{G}_g \underset{\sim}{\Sigma}^{-1} \underset{\sim}{G}_h y_g y_h$$

$$\leq \sum_{g,h=0}^{m} \operatorname{tr} \underset{\sim}{\Sigma}_L^{-1} \underset{\sim}{G}_g \underset{\sim}{\Sigma}_L^{-1} \underset{\sim}{G}_h y_g y_h$$

for every $(y_0, y_1, \ldots , y_m)$. Then for $\underset{\sim}{B}$ defined by (87) of Section 10.2 of T. W. Anderson (1971) with T replaced by p ,

$$\text{(3.73)} \qquad \sum_{g,h=0}^{m} \operatorname{tr} \underset{\sim}{\Sigma}_U^{-1} \underset{\sim}{C}_g \underset{\sim}{\Sigma}_U^{-1} \underset{\sim}{G}_h y_g y_h = \sum_{g,h=0}^{m} \operatorname{tr} \underset{\sim}{B}'\underset{\sim}{B} \underset{\sim}{G}_g \underset{\sim}{B}'\underset{\sim}{B} \underset{\sim}{G}_h y_g y_h$$

$$= \sum_{g,h=0}^{m} \sum_{\substack{i,j,k,\ell,\\ m,n=1}}^{p} b_{ji} b_{jk} g_{k\ell}^{(g)} b_{q\ell} b_{qn} g_{nj}^{(h)} y_g y_h$$

$$= \sum_{j,q=1}^{p} \left(\sum_{g=0}^{m} \sum_{k, =1}^{p} b_{jk} g_{k_\ell}^{(g)} b_{q_\ell} y_g \right)^2$$

$$\geq \sum_{j,q=K+1}^{p} \left(\sum_{g=0}^{m} \sum_{k,\ell=1}^{p} b_{jk} g_{k}^{(g)} b_{q} y_g \right)^2$$

$$= \sum_{j,q=K+1}^{p} \left(\sum_{g=0}^{m} \sum_{k',\ell'=0}^{K} b_{k'} b_{\ell'} g_{j-k'q-\ell'}^{(g)} y_g \right)^2 ,$$

where K is the order of the approximating autoregressive process.

We have

$$\text{(3.74)} \qquad g_{ab}^{(0)} = 1 , \qquad a = b ,$$

$$= 0 , \qquad a \neq b ,$$

$$(3.75)\qquad g_{ab}^{(h)} = 1\,, \qquad a = b \pm h,\ h > 0\,,$$
$$= 0\,, \qquad a \neq b \pm h,\ h > 0\,,$$

where a, b, and h are integers. Then

$$(3.76)\qquad g_{ab}^{(0)} = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i\lambda(a-b)}\, d\lambda\,,$$

$$(3.77)\qquad g_{ab}^{(h)} = \frac{1}{2\pi} \int_{-\pi}^{\pi} [e^{i\lambda(a-b+h)} + e^{i\lambda(a-b-h)}]d\lambda\,, \qquad h > 0\,.$$

If $\delta_h = \frac{1}{2}$ for $h=0$ and $\delta_h = 1$ for $h > 0$, then

$$(3.78)\qquad g_{ab}^{(h)} = \frac{\delta_h}{2\pi} \int_{-\pi}^{\pi} [e^{i\lambda(a-b+h)} + e^{i\lambda(a-b-h)}]d\lambda\,, \quad h=0, 1, \ldots, m\,.$$

Then $1/(p-K)$ times the right-hand side of (3.73) is

$$(3.79)\qquad \frac{1}{p-K} \sum_{j,q=K+1}^{p} \left(\sum_{g=0}^{m} \sum_{k',\ell'=0}^{K} b_{k'} b_{\ell'} y_g \right.$$
$$\left. \frac{\delta_g}{2\pi} \int_{-\pi}^{\pi} \left\{ e^{i\lambda[(j-k')-(q-\ell')+g]} + e^{i\lambda[(j-k')-(q-\ell')-g]} \right\} d\lambda \right)^2$$
$$= \frac{1}{2\pi(p-K)} \sum_{j,q=K+1}^{p} \left(\int_{-\pi}^{\pi} e^{i\lambda(j-q)} \sum_{k',\ell'=0}^{K} b_{k'} e^{-i\lambda k'} \right.$$
$$\left. b_{\ell'} e^{i\lambda\ell'} \delta_g \left\{ e^{i\lambda g} + e^{-i\lambda g} \right\} d\lambda \right)^2$$
$$= \frac{1}{(2\pi)(p-K)} \sum_{j,q=K+1}^{p} \left(2 \sum_{g=0}^{m} y_g \delta_g \int_{-\pi}^{\pi} e^{i\lambda(j-q)} \frac{\cos \lambda g}{2\pi\, f_U(\lambda)}\, d\lambda \right)^2$$

because

$$(3.80)\qquad f_U(\lambda) = \frac{1}{2\pi \left| \sum_{k=0}^{K} b_k e^{i\lambda k} \right|^2}\,.$$

Then the limit of (3.79) is

$$(3.81)\qquad \frac{4}{(2\pi)^2}\sum_{g,h=0}^{m} y_g\, y_h\, \delta_g\, \delta_h \lim_{p\to\infty}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\sum_{j,q=1}^{p-K}\frac{1}{2\pi(p-K)}\, e^{i(\lambda+\nu)(j-q)}$$

$$\frac{\cos \lambda g \cos \nu h}{f_U(\lambda)\; f_U(\nu)}\, d\lambda\, d\nu$$

$$= \frac{4}{(2\pi)^2}\sum_{g,h=0}^{m} y_g\, y_h\, \delta_g\, \delta_h \lim_{p\to\infty}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} k_{p-K}\,(\lambda+\nu)$$

$$\frac{\cos \lambda g \cos \nu h}{f_U(\lambda)\; f_U(\nu)}\, d\lambda\, d\nu\ ,$$

where

$$(3.82)\qquad k_{p-K}\,(\lambda+\nu) = \frac{\sin^2 \frac{1}{2}\,(\lambda+\nu)(p-K)}{2\pi(p-K)\,\sin^2 \frac{1}{2}\,(\lambda+\nu)}\ .$$

Then (3.81) is

$$(3.83)\qquad \frac{4}{(2\pi)^2}\sum_{g,h=0}^{m} y_g\, y_h\, \delta_g\, \delta_h \lim_{p\to\infty}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} k_{p-K}\,(\lambda-\mu)\,\frac{\cos \lambda g \cos \mu h}{f_U(\lambda)\; f_U(\mu)}\, d\lambda\, d\mu$$

$$= \frac{4}{(2\pi)^2}\sum_{g,h=0}^{m} y_g\, y_h\, \delta_g\, \delta_h \int_{-\pi}^{\pi}\frac{\cos \lambda g \cos \lambda h}{f_U^2\,(\lambda)}\, d\lambda$$

by substitution $\nu = -\mu$ and the argument leading to Theorem 8.3.3 of T. W. Anderson (1971). Then

$$(3.84)\qquad \lim_{p\to\infty}\frac{1}{p}\sum_{g,h=0}^{m} y_g\, y_h\, \mathrm{tr}\ \underset{\sim}{\Sigma}^{-1}\, \underset{\sim}{G}_g\, \underset{\sim}{\Sigma}^{-1}\, \underset{\sim}{G}_h$$

is greater than or equal to (3.83). Similarly

$$(3.85)\qquad \overline{\lim_{p\to\infty}} \frac{1}{p} \sum_{g,h=0}^{m} y_g\, y_h \operatorname{tr} \underset{\sim}{\Sigma}^{-1} \underset{\sim}{G}_g \underset{\sim}{\Sigma}^{-1} \underset{\sim}{G}_h$$

$$\leq \frac{4}{2\pi} \sum_{g,h=0}^{m} y_g\, y_h\, \delta_g\, \delta_h \int_{-\pi}^{\pi} \frac{\cos \lambda g \cos \lambda h}{f_L^2(\lambda)}\, d\lambda\ .$$

Since ε is arbitrary

$$(3.86)\qquad \lim_{p\to\infty} \frac{1}{p} \sum_{g,h=0}^{m} y_g\, y_h \operatorname{tr} \underset{\sim}{\Sigma}^{-1} \underset{\sim}{G}_g \underset{\sim}{\Sigma}^{-1} \underset{\sim}{G}_h$$

$$= \frac{4}{(2\pi)^2} \sum_{g,h=0}^{m} y_g\, y_h\, \delta_g\, \delta_h \int_{-\pi}^{\pi} \frac{\cos \lambda g \cos \lambda h}{f^2(\lambda)}\, d\lambda\ .$$

Since (3.86) holds for every set of $y_0, y_1, \ldots, y_m$, the theorem follows.

It is of interest to compare this result with Theorem 8.3.3 of T. W. Anderson (1971) which gives the asymptotic covariances of the sample covariances of a stationary process with a continuous spectral density. The spectral density of an autoregressive process with coefficients $\alpha_1, \ldots, \alpha_m$ is proportional to the reciprocal of $f(\lambda)$ for the moving average process; hence the Gaussian part of (37) of Section 8.3 of T. W. Anderson (1971) is proportional to the coefficient of $y_g\, y_h$ in (3.86).

4. Asymptotic Theory as the Sample Size Increases

When $\underset{\sim}{x}_1, \ldots, \underset{\sim}{x}_N$ are observed with $\mathcal{E}\underset{\sim}{x} = \underset{\sim}{\mu}$ satisfying (1.2) and $\underset{\sim}{\Sigma}$ is known, the Markov estimates of $\beta_1, \ldots, \beta_r$ are the solutions to

$$\sum_{i=1}^{r} \underset{\sim}{z}_j' \underset{\sim}{\Sigma}^{-1} \underset{\sim}{z}_i \hat{\beta}_i = \underset{\sim}{z}_j' \underset{\sim}{\Sigma}^{-1} \bar{\underset{\sim}{x}}_N , \qquad j=1, \ldots, r . \tag{4.1}$$

In matrix notation the solution is

$$\hat{\underset{\sim}{\beta}}_N = (\underset{\sim}{Z}' \underset{\sim}{\Sigma}^{-1} \underset{\sim}{Z})^{-1} \underset{\sim}{Z}' \underset{\sim}{\Sigma}^{-1} \bar{\underset{\sim}{x}}_N , \tag{4.2}$$

where $\hat{\underset{\sim}{\beta}}_N = (\hat{\beta}_1, \ldots, \hat{\beta}_r)'$ and

$$\underset{\sim}{Z} = (\underset{\sim}{z}_1, \ldots, \underset{\sim}{z}_r) . \tag{4.3}$$

(We use the subscript N on $\bar{\underset{\sim}{x}}_N$ and $\hat{\underset{\sim}{\beta}}_N$ to emphasize the dependence on N .) Then the covariance of $\sqrt{N}\,(\hat{\underset{\sim}{\beta}}_N-\underset{\sim}{\beta})$ is

$$\mathcal{C}[\sqrt{N}\,(\hat{\underset{\sim}{\beta}}_N-\underset{\sim}{\beta})] = N\,\mathcal{E}(\hat{\underset{\sim}{\beta}}_N-\underset{\sim}{\beta})(\hat{\underset{\sim}{\beta}}_N-\underset{\sim}{\beta})' = (\underset{\sim}{Z}' \underset{\sim}{\Sigma}^{-1} \underset{\sim}{Z})^{-1} . \tag{4.4}$$

Regardless of the distribution of $\underset{\sim}{X}$, $\sqrt{N}\,(\hat{\underset{\sim}{\beta}}_N-\underset{\sim}{\beta})$ has a limiting normal distribution with mean vector $\underset{\sim}{0}$ and covariance matrix (4.4), because $\sqrt{N}\,(\bar{\underset{\sim}{x}}_N-\underset{\sim}{\mu})$ has a limiting normal distribution.

If $\underset{\sim}{\Sigma}$ is unknown, let us suppose that we have a consistent estimate $\hat{\underset{\sim}{\Sigma}}_N$ of $\underset{\sim}{\Sigma}$. Consider the estimate

$$\hat{\underset{\sim}{\beta}}{}^*_N = (\underset{\sim}{Z}' \hat{\underset{\sim}{\Sigma}}{}_N^{-1} \underset{\sim}{Z})^{-1} \underset{\sim}{Z}' \hat{\underset{\sim}{\Sigma}}{}_N^{-1} \bar{\underset{\sim}{x}}_N . \tag{4.5}$$

Then

$$(4.6) \qquad \sqrt{N}\,(\hat{\beta}^*_N - \hat{\beta}) = \sqrt{N}\,[\hat{\beta}^*_N - \beta - (\hat{\beta}_N - \beta)]$$

$$= [(Z'\,\hat{\Sigma}_N^{-1}\,Z)^{-1}\,Z'\,\hat{\Sigma}_N^{-1} - (Z'\,\Sigma^{-1}\,Z)^{-1}\,Z'\,\Sigma^{-1}]\,\sqrt{N}\,(\bar{x}_N - Z\beta)$$

converges stochastically to 0 because

$$(4.7) \qquad \operatorname*{plim}_{N\to\infty} (Z'\,\hat{\Sigma}_N^{-1}\,Z)^{-1}\,Z'\,\hat{\Sigma}_N^{-1} = (Z'\,\Sigma^{-1}\,Z)^{-1}\,Z'\,\Sigma^{-1}$$

and $\sqrt{N}\,(\bar{x}_N - Z\beta) = \sqrt{N}\,(\bar{x}_N - \mu)$ has a limiting distribution. Thus $\sqrt{N}\,(\hat{\beta}^*_N - \beta)$ has a limiting normal distribution with mean 0 and covariance matrix $(Z'\,\Sigma^{-1}\,Z)^{-1}$, which is the same as the limiting normal distribution of $\hat{\beta}_N$. If $\hat{\beta}_N$ is asymptotically efficient, then $\hat{\beta}^*_N$ is (in the same sense). In particular, when $\hat{\beta}_N$ is maximum likelihood, as when X has a normal distribution, it is asymptotically efficient in the sense of attaining the Cramér-Rao lower bound for the covariance matrix of unbiased estimates. We summarize the result.

Theorem 4.1. Let $x_1, \ldots, x_N$ be identically distributed with mean $\mathcal{E}X = Z\beta$ and covariance matrix Σ and let $\hat{\Sigma}_N$ be a consistent estimate of Σ. Then, if $\hat{\beta}^*_N$ is given by (4.5), $\sqrt{N}\,(\hat{\beta}^*_N - \beta)$ has a limiting normal distribution with covariance matrix $(Z\,\Sigma^{-1}\,Z)^{-1}$. If $\hat{\beta}_N$ given by (4.2) is asymptotically efficient, so is $\hat{\beta}^*_N$.

We now apply the result to the estimation of $\sigma_0, \sigma_1, \ldots, \sigma_m$.

Theorem 4.2. Let $x_1, \ldots, x_N$ be N observations from $N(\mu, \Sigma)$, where μ is known and Σ is given by (1.3). Let C be defined by (1.7). Let $\hat{\sigma}_0^{(0)}, \hat{\sigma}_1^{(0)}, \ldots, \hat{\sigma}_m^{(0)}$ be a consistent set of estimates of $\sigma_0, \sigma_1, \ldots, \sigma_m$. Let $\hat{\sigma}_0^{(1)}, \hat{\sigma}_1^{(1)}, \ldots, \hat{\sigma}_m^{(1)}$ be the solution

to (2.1) for $i=1$. Then $\sqrt{N}\,(\hat{\sigma}_0^{(1)}-\sigma_0)$, $\sqrt{N}\,(\hat{\sigma}_1^{(1)}-\sigma_1)$, ... , $\sqrt{N}\,(\hat{\sigma}_m^{(1)}-\sigma_m)$ have a limiting normal distribution with means 0 and covariance matrix (1.10), and $\hat{\sigma}_0^{(1)}$, $\hat{\sigma}_1^{(1)}$, ... , $\hat{\sigma}_m^{(1)}$ are asymptotically efficient.

Proof. In Theorem 1 replace $\bar{\underset{\sim}{x}}_N$ by $\underset{\sim}{C}$, $\underset{\sim}{Z}\underset{\sim}{\beta}$ by $\sum_{h=0}^{m} \sigma_h \underset{\sim}{g}^{(h)}$, $\underset{\sim}{\Sigma}$ by $\underset{\sim}{\Phi}$, and $\hat{\underset{\sim}{\Sigma}}_N$ by $\hat{\underset{\sim}{\Phi}}_N$ inserting $\hat{\sigma}_0^{(0)}$, $\hat{\sigma}_1^{(0)}$, ... , $\hat{\sigma}_m^{(0)}$ in $\underset{\sim}{\Phi}$. Then $\hat{\underset{\sim}{\Phi}}_N$ is a consistent estimate of $\underset{\sim}{\Phi}$ and the limiting normal distribution follows. Since the solutions to (1.22) are asymptotically efficient, the theorem follows.

If $\underset{\sim}{\mu}$ is unknown, it can be estimated by $\bar{\underset{\sim}{x}}_N$, and $\underset{\sim}{C}$ in Theorem 2 can be replaced by $(1/N)\sum_{\alpha=1}^{N} (\underset{\sim}{x}_\alpha-\bar{\underset{\sim}{x}}_N)(\underset{\sim}{x}_\alpha-\bar{\underset{\sim}{x}}_N)'$. If $\underset{\sim}{\mu}$ has the form (1.2), it can be estimated by $\sum_{j=1}^{r} \hat{\beta}_j \underset{\sim}{z}_j$, where $\hat{\beta}_1$, ... , $\hat{\beta}_r$ is a solution to (2.21). For the asymptotic theory uses the fact that $\sqrt{N}\,(\underset{\sim}{C}-\underset{\sim}{\Sigma})$ has a limiting normal distribution,

$$\sqrt{N}\,[\underset{\sim}{C} - \frac{1}{N}\sum_{\alpha=1}^{N} (\underset{\sim}{x}_\alpha-\bar{\underset{\sim}{x}})(\underset{\sim}{x}_\alpha-\bar{\underset{\sim}{x}})'] = \sqrt{N}\,(\bar{\underset{\sim}{x}}_N-\underset{\sim}{\mu})(\bar{\underset{\sim}{x}}_N-\underset{\sim}{\mu})', \tag{4.8}$$

$$\sqrt{N}\,[\underset{\sim}{C}-\hat{\underset{\sim}{C}}] = \sqrt{N}\,(\bar{\underset{\sim}{x}}_N-\underset{\sim}{\mu})(\bar{\underset{\sim}{x}}_N-\underset{\sim}{\mu})' - \sqrt{N}\,(\bar{\underset{\sim}{x}}_N-\hat{\underset{\sim}{\mu}})(\bar{\underset{\sim}{x}}_N-\hat{\underset{\sim}{\mu}})' \tag{4.9}$$

and the facts that $\operatorname{plim}_{N\to\infty} \bar{\underset{\sim}{x}}_N = \underset{\sim}{\mu}$, $\operatorname{plim}_{N\to\infty} \hat{\underset{\sim}{\mu}} = \underset{\sim}{\mu}$, and $\sqrt{N}\,(\bar{\underset{\sim}{x}}_N-\underset{\sim}{\mu})$ and $\sqrt{N}\,(\bar{\underset{\sim}{x}}_N-\hat{\underset{\sim}{\mu}})$ have limiting (normal) distributions.

5. Estimation of the Coefficients of a Moving Average Process

We now treat the problem of estimating the coefficients $\alpha_0, \alpha_1, \ldots, \alpha_m$ of the moving average process (1.25) with the restriction $\sigma^2 = 1$ replacing the restriction $\alpha_0 = 1$. Then

$$\sigma_h = \sum_{g=0}^{m-h} \alpha_g \alpha_{g+h}, \qquad h=0, 1, \ldots, m. \tag{5.1}$$

Let

$$M(z) = \sum_{j=0}^{m} \alpha_j z^{m-j}. \tag{5.2}$$

Then

$$\sum_{h=-m}^{m} \sigma_h z^h = M(z) M(z^{-1}) \tag{5.3}$$

by (5.1).

Let the roots of $M(z) = 0$, which is (1.32), be $z_1, \ldots, z_m$. Then

$$M(z) = \alpha_0 \prod_{j=1}^{m} (z-z_j) \tag{5.4}$$

and

$$\begin{aligned}\sum_{h=-m}^{m} \sigma_h z^{m+h} &= \alpha_0^2 z^m \left[\prod_{j=1}^{m} (z-z_j)\right]\left[\prod_{j=1}^{m} \left(\frac{1}{z} - z_j\right)\right] \\ &= \alpha_0^2 \left[\prod_{j=1}^{m} (z-z_j)\right]\left[\prod_{j=1}^{m} (1-z\, z_j)\right].\end{aligned} \tag{5.5}$$

Thus, if $\sigma(m) \neq 0$, the $2m$ roots of

$$\sum_{h=-m}^{m} \sigma_h z^{m+h} = 0 \tag{5.6}$$

are $z_1, \ldots, z_m, 1/z_1, \ldots, 1/z_m$.

Conversely, if $\sigma_0, \sigma_1, \ldots, \sigma_m \neq 0$ are given, the set of $2m$ roots of (5.6) can be partitioned into two sets of m roots each, such that the roots in one set are less than or equal to 1 in absolute value and are the roots of an equation (1.32) with real coefficients, that is, $M(z) = 0$, defined by (5.2) with real coefficients, and the roots in the other set are reciprocals of the respective roots in the first set. The roots define the coefficients in $M(z)$ except for normalization, which is determined by (5.1) for $k = 0$. [See Section 5.7 of T. W. Anderson (1971).]

The equation (5.6) can be written

(5.7)
$$0 = \sigma_0 z^m + \sum_{h=1}^{m} \sigma_h (z^{m+h} + z^{m-h})$$
$$= z^m \left[\sigma_0 + \sum_{h=1}^{m} \sigma_h \frac{z^{2h} + 1}{z^h} \right] .$$

Let

(5.8)
$$w = z + \frac{1}{z} = \frac{z^2 + 1}{z} ;$$

that is,

(5.9)
$$zw = z^2 + 1 .$$

Then

(5.10)
$$z^{2r} w^{2r} = (zw)^{2r} = (z^2+1)^{2r}$$
$$= \left[(z^2)^{2r}+1\right] + 2r\, z^2 \left[(z^2)^{2r-2} + 1\right] + \binom{2r}{2} (z^2)^2 \left[(z^2)^{2r-4} + 1\right]$$
$$+ \ldots + \binom{2r}{r-1} (z^2)^{r-1} \left[(z^2)^2 + 1\right] + \binom{2r}{r} z^{2r} ,$$

$$(5.11)\quad z^{2r+1} w^{2r+1} = (zw)^{2r+1} = (z^2+1)^{2r+1}$$

$$= \left[(z^2)^{2r+1} + 1\right] + (2r+1)\, z^2 \left[(z^2)^{2r-2} + 1\right]$$

$$+ \binom{2r+1}{2} (z^2)^2 \left[(z^2)^{2r-4} + 1\right] + \dots + \binom{2r+1}{r} (z^2)^r \left[z^2+1\right] .$$

[The coefficients of $(z^2)^s + 1$ in (5.10) and (5.11) are the same as the coefficients of A_s in (16) and (17) of Section 6.5 of T. W. Anderson (1971).] When we solve (5.10) and (5.11) successively for $(z^2)^s + 1$, s=1, 2, ... , m, we obtain

$$(5.12)\quad (z^2)^{2r}+1 = z^{2r} \left[w^{2r} + c_{2r,2}\, w^{2r-2} + \dots + c_{2r,2r}\right] ,$$

$$(5.13)\quad (z^2)^{2r+1} + 1 = z^{2r+1} \left[w^{2r+1} + c_{2r+1,2}\, w^{2r-1} + \dots + c_{2r+1,2r}\, w\right] ,$$

where $c_{2r,2}, \dots , c_{2r,2r}, c_{2r+1,2}, \dots , c_{2r+1,2r}$ are appropriate constants. Then (5.7) can be written

$$(5.14)\quad 0 = z^{2n} \{\sigma_0 + \sigma_1 w + \sigma_2 [w^2-2]$$

$$+ \dots + \sigma_{2n} [w^{2n} + c_{2n,2}\, w^{2n-2} + \dots + c_{2n,2n}]\}$$

$$= z^{2n} \{\sigma_{2n}\, w^{2n} + \sigma_{2n-1}\, w^{2n-1} + \dots + \sigma_0 - 2\sigma_2 + \dots + \sigma_{2n}\, c_{2n,2n}\} ,$$

$m=2n$

$$(5.15)\quad 0 = z^{2n+1} \{\sigma_0 + \sigma_1 w + \sigma_2 [w^2-2]$$

$$+ \dots + \sigma_{2n+1} [w^{2n+1} + c_{2n+1,2}\, w^{2n-1} + \dots + c_{2n+1,2n}\, w]\}$$

$$= z^{2n+1} \{\sigma_{2n+1}\, w^{2n+1} + \sigma_{2n}\, w^{2n} + \dots + \sigma_0 - 2\sigma_2 + \dots + \sigma_{2n}\, c_{2n,2n}\} ,$$

$m = 2n+1$.

Solving (5.14) or (5.15) as an equation in w, we obtain m roots, $w_1, \ldots, w_m$. Then solve

$$z^2 - w_i z + 1 = 0 \tag{5.16}$$

for z, to obtain the roots

$$\frac{w_i}{2} \pm \sqrt{\left(\frac{w_i}{2}\right)^2 - 1} \tag{5.17}$$

to (5.7); the pair of roots to (5.17) are reciprocals. The m roots with absolute value less than or equal to 1 are the desired roots of $M(z) = 0$ and the coefficients of the polynomial $M(z)$ are the desired $\alpha_0, \alpha_1, \ldots, \alpha_m$ except for a constant of proportionality.

We want to modify the numerical procedures discussed in Section 2. If $\hat{\sigma}_0^{(i-1)}, \hat{\sigma}_1^{(i-1)}, \ldots, \hat{\sigma}_m^{(i-1)}$ are the estimates at the i-th stage, let $\hat{\alpha}_0^{(i-1)}, \hat{\alpha}_1^{(i-1)}, \ldots, \hat{\alpha}_m^{(i-1)}$ be the desired solution of (5.1) with the σ_h's replaced by the $\hat{\sigma}_h^{(i-1)}$'s. We want to determine the next iteration; let

$$\hat{\alpha}_h^{(i)} = \hat{\alpha}_h^{(i-1)} + d_h, \qquad h=0, 1, \ldots, m . \tag{5.18}$$

Then if we substitute into (2.1) with $\underset{\sim}{C} = \underset{\sim}{x}\underset{\sim}{x}'$ we obtain

$$\sum_{f=0}^{m} \operatorname{tr} \underset{\sim}{\hat{\Sigma}}{}_{i-1}^{-1} \underset{\sim}{G}_g \underset{\sim}{\hat{\Sigma}}{}_{i-1}^{-1} \underset{\sim}{G}_f \sum_{h=0}^{m-f} (\hat{\alpha}_h^{(i-1)} + d_n)(\hat{\alpha}_{h+f}^{(i-1)} + d_{h+f}) \tag{5.19}$$

$$= \underset{\sim}{x}' \underset{\sim}{\hat{\Sigma}}{}_{i-1}^{-1} \underset{\sim}{G}_g \underset{\sim}{\hat{\Sigma}}{}_{i-1}^{-1} \underset{\sim}{x} , \qquad g=0, 1, \ldots, m .$$

The left-hand side of (5.19) can be written

$$(5.20)\quad \sum_{f=0}^{m} \operatorname{tr} \hat{\Sigma}_{i-1}^{-1} G_g \hat{\Sigma}_{i-1}^{-1} G_f \left[\hat{\sigma}_f^{(i-1)} + \sum_{h=0}^{m-f} (\hat{\alpha}_h^{(i-1)} d_{h+f} + \hat{\alpha}_{h+f}^{(i-1)} d_h + d_{h+f} d_h)\right]$$

$$= \operatorname{tr} \hat{\Sigma}_{i-1}^{-1} G_g + \sum_{j=0}^{m} \left[\sum_{k=0}^{j} \operatorname{tr} \hat{\Sigma}_{i-1}^{-1} G_g \hat{\Sigma}_{i-1}^{-1} G_k \hat{\alpha}_{j-k}^{(i-1)} + \sum_{k=0}^{m-j} \operatorname{tr} \hat{\Sigma}_{i-1}^{-1} G_g \hat{\Sigma}_{i-1}^{-1} G_k \hat{\alpha}_{j+k}^{(i-1)} \right] d_j + \sum_{f=0}^{m} \operatorname{tr} \hat{\Sigma}_{i-1}^{-1} G_g \hat{\Sigma}_{i-1}^{-1} G_f \sum_{h=0}^{m-f} d_h d_{h+f} .$$

Using the linear terms only, we can obtain from (5.19) the equations

$$(5.21)\quad \sum_{j=0}^{m} \left[\sum_{k=0}^{j} \operatorname{tr} \hat{\Sigma}_{i-1}^{-1} G_g \hat{\Sigma}_{i-1}^{-1} G_k \hat{\alpha}_{j-k}^{(i-1)} + \sum_{k=0}^{m-j} \operatorname{tr} \hat{\Sigma}_{i-1}^{-1} G_g \hat{\Sigma}_{i-1}^{-1} G_k \hat{\alpha}_{j+k}^{(i-1)} \right] d_j$$

$$= x' \hat{\Sigma}_{i-1}^{-1} G_g \hat{\Sigma}_{i-1}^{-1} x - \operatorname{tr} \hat{\Sigma}_{i-1}^{-1} G_g , \qquad g=0, 1, \ldots , m .$$

If $\hat{\sigma}_0^{(0)}, \hat{\sigma}_1^{(0)}, \ldots , \hat{\sigma}_m^{(0)}$ are consistent estimates of $\sigma_0, \sigma_1, \ldots , \sigma_m$, then $\hat{\alpha}_0^{(0)}, \hat{\alpha}_1^{(0)}, \ldots , \hat{\alpha}_m^{(0)}$ are consistent estimates of $\alpha_0, \alpha_1, \ldots , \alpha_m$. If the solution to (5.21) for $i = 0$ is $d_0, d_1, \ldots , d_m$, then $\hat{\alpha}_0^{(0)} + d_0$, $\hat{\alpha}_1^{(0)} + d_1, \ldots , \hat{\alpha}_m^{(0)} + d_m$ are consistent estimates of $\alpha_0, \alpha_1, \ldots , \alpha_m$. It may be expected that these estimates are asymptotically efficient as $p \to \infty$.

A method suggested by Durbin (1959) for estimating $\alpha_1, \ldots , \alpha_m$ when $\alpha_0 = 1$ is to solve

$$(5.22)\quad \sum_{j=0}^{n-1} c_{i-j} \hat{\gamma}_j = - c_i , \qquad i=1, \ldots , n$$

for some $n > m$, where

$$(5.23) \qquad c_j = c_{-j} = \frac{1}{p} \sum_{i=1}^{p-j} x_i x_{i+j} , \qquad j=0, 1, \ldots , n ,$$

and let $\hat{\gamma}_0 = 1$. Then solve

$$(5.24) \qquad \sum_{g=1}^{m} h_{f-g} \hat{\alpha}_g = -h_f , \qquad f=1, \ldots , m ,$$

where

$$(5.25) \qquad h_g = h_{-g} = \sum_{f=1}^{n-g} \hat{\gamma}_f \hat{\gamma}_{f+g} , \qquad g=1, \ldots , m .$$

See T. W. Anderson (1971), Section 5.7.2, for more explanation and discussion. Raul Mentz has shown that the resulting estimates are approximately consistent and asymptotically normally distributed as $p \to \infty$. These estimates (suitably normalized) could be used as starting values in the above iteration procedure.

Since

$$(5.26) \qquad \operatorname*{plim}_{p\to\infty} c_j = \hat{\gamma}_j , \qquad j=0, 1, \ldots , m ,$$

$$(5.27) \qquad \operatorname*{plim}_{p\to\infty} c_j = 0 , \qquad j=m+1, \ldots , n ,$$

the equations (5.22) can be replaced by

$$(5.28) \qquad \underset{\sim}{C}^* \underset{\sim}{\hat{\gamma}} = \underset{\sim}{c} ,$$

where $\underset{\sim}{C}^*$ has the form of (1.29) with σ_g replaced by c_g, $g=0, 1, \ldots , m$, $\underset{\sim}{\hat{\gamma}} = (\hat{\gamma}_1, \ldots , \hat{\gamma}_n)'$, and $\underset{\sim}{c} = (c_1, \ldots , c_n)'$. The forward solution can be applied to (5.28) and the forms $h_0, \ldots , h_m$ can be formed as indicated in Section 2 with p replaced by n.

REFERENCES

Anderson, T. W. (1958), An Introduction to Multivariate Statistical Analysis, John Wiley & Sons, Inc., New York.

Anderson, T. W. (1969), "Statistical inference for covariance matrices with linear structure", Proceedings of the Second International Symposium on Multivariate Analysis (P. R. Krishnaiah, ed.), Academic Press, Inc., New York, 1969, pp. 55-66.

Anderson, T. W. (1970), "Estimation of covariance matrices which are linear combinations or whose inverses are linear combinations of given matrices", Essays in Probability and Statistics, University of North Carolina Press, Chapel Hill, 1970, pp. 1-24. Also Technical Report No. 3, 1966, Contract AF 41(609)-2653, Teacher's College, Columbia University.

Anderson, T. W. (1971), The Statistical Analysis of Time Series, John Wiley & Sons, Inc., New York.

Durbin, J. (1959), "Efficient estimation of parameters in moving-average models", J. Roy. Statist. Soc. Ser. B, **22**, pp. 139-153.

Hartley, H. O. and J. N. K. Rao, "Maximum likelihood estimator for the mixed analysis of variance model", Biometrika, **54**, pp. 93-108.

Shaman, P. (1969), "On the inverse of the covariance matrix of a first order moving average", Biometrika, **56**, pp. 595-608.

UNCLASSIFIED
Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) DEPARTMENT OF STATISTICS STANFORD UNIVERSITY CALIFORNIA	2a. REPORT SECURITY CLASSIFICATION 2b. GROUP

3. REPORT TITLE
ESTIMATION OF COVARIANCE MATRICES WITH LINEAR STRUCTURE AND MOVING AVERAGE PROCESSES OF FINITE ORDER

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)
TECHNICAL REPORT

5. AUTHOR(S) (Last name, first name, initial)
ANDERSON, T.W.

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
OCTOBER 29, 1971	48	7

8a. CONTRACT OR GRANT NO. N00014-67-A-0112-0030 b. PROJECT NO. NR-042-034 c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) #6 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. AVAILABILITY/LIMITATION NOTICES
Reproduction in whole or in part is permitted for any purpose of the United States Government

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	OFFICE OF NAVAL RESEARCH ARLINGTON, VIRGINIA

13. ABSTRACT

One or more observations are made on a random vector, whose covariance matrix may be a linear combination of known symmetric matrices and whose mean vector may be a linear combination of known vectors; the coefficients of the linear combinations are unknown parameters to be estimated. Under the assumption of normality equations are developed for the maximum likelihood estimates. The solution of these equations by iterative methods is indicated. A sequence of observations on a moving-average process of finite order may be considered as an observation on a random vector whose covariance matrix has the linear structure. The general method of estimation applied to this problem is particularly easy to compute.

DD FORM 1 JAN 64 1473

UNCLASSIFIED
Security Classification

UNCLASSIFIED

Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
maximum likelihood estimates						
covariance matrices						
multivariate normal distribution						
moving average process						
asymptotically efficient estimates						
iterative computations						

INSTRUCTIONS

1. ORIGINATING ACTIVITY: Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization (*corporate author*) issuing the report.

2a. REPORT SECURITY CLASSIFICATION: Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2b. GROUP: Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. REPORT TITLE: Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. DESCRIPTIVE NOTES: If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. AUTHOR(S): Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. REPORT DATE: Enter the date of the report as day, month, year; or month, year. If more than one date appears on the report, use date of publication.

7a. TOTAL NUMBER OF PAGES: The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7b. NUMBER OF REFERENCES: Enter the total number of references cited in the report.

8a. CONTRACT OR GRANT NUMBER: If appropriate, enter the applicable number of the contract or grant under which the report was written.

8b, 8c, & 8d. PROJECT NUMBER: Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9a. ORIGINATOR'S REPORT NUMBER(S): Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9b. OTHER REPORT NUMBER(S): If the report has been assigned any other report numbers (*either by the originator or by the sponsor*), also enter this number(s).

10. AVAILABILITY/LIMITATION NOTICES: Enter any limitations on further dissemination of the report, other than those imposed by security classification, using standard statements such as:

(1) "Qualified requesters may obtain copies of this report from DDC."

(2) "Foreign announcement and dissemination of this report by DDC is not authorized."

(3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through ________________."

(4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through ________________."

(5) "All distribution of this report is controlled. Qualified DDC users shall request through ________________."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. SUPPLEMENTARY NOTES: Use for additional explanatory notes.

12. SPONSORING MILITARY ACTIVITY: Enter the name of the departmental project office or laboratory sponsoring (*paying for*) the research and development. Include address.

13. ABSTRACT: Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as *(TS), (S), (C), or (U)*.

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. KEY WORDS: Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, roles, and weights is optional.

DD FORM 1 JAN 64 1473 (BACK)

Unclassified

Security Classification